The Effect of Time Pressure on Map-Based Decision Making

Dissertation
zur
Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der
Mathematisch-naturwissenschaftlichen Fakultät
der
Universität Zürich

von

Jan Wilkening
aus
Deutschland

Promotionskomitee

Prof. Dr. Sara Irina Fabrikant (Vorsitz)
Dr. Tumasch Reichenbacher
Prof. Dr. Daniel R. Montello

Zürich, 2012

ABSTRACT

In our everyday lives, many decisions are made based on different maps and under different time pressure situations. Some map-based decisions under time pressure can even decide over life and death. Hence, it is important to understand how different time constraints and map types influence the quality of map-based decisions, and which role the characteristics of the decision-makers play in this context.

While several empirical map use studies have investigated the efficiency and effectiveness of decisions with different maps, very little is known about the relevance of time pressure in this regard. This thesis bridges this research gap with four controlled user experiments on map-based decision making under time pressure with laymen. These experiments are complemented by several expert interviews with people who make map-based decisions under time pressure in their daily lives.

For all the four experiments, time pressure is operationalized as an independent, exogenous factor. First, I assess how time pressure influences the preference for map display types and interaction tools in road selection tasks (Experiment I). Then, I investigate how time limits affect response accuracy and confidence in a similar task with road maps and satellite images (Experiment II). The next experiment focuses on a slope detection task with 2D and 3D maps (Experiment III). Finally, I explore the effect of time pressure on human-map interactions and on decisions humans make in different map-based tasks with a highly interactive virtual globe (Experiment IV).

Overall, the results suggest that time pressure indeed influences map-based decisions in several ways. In the road selection task, different time limits affect response confidence more than response accuracy. In the slope detection task, both response accuracy and confidence decrease significantly with more decision time available. Furthermore, time pressure influences the frequency of human-map interactions and the preference for map display types and interaction tools.

Moreover, the empirical results indicate that more abstract 2D maps with less visual clutter might be more suitable for map-based decisions under time pressure than realistic 3D displays, but also that people prefer more familiar map types in this context. Overall, this study contributes to a better understanding of how humans use maps in time-critical situations, and thereby provides first insights on how maps should be designed for effective and efficient spatio-temporal inference and decision making in such situations.

ZUSAMMENFASSUNG

Viele Entscheidungen in unserem täglichen Leben werden mit verschiedenen Karten und unter verschiedenen Zeitdrucksituationen getroffen. Manche dieser kartenbasierten Entscheidungen unter Zeitdruck können sogar zwischen Leben und Tod entscheiden. Daher ist es wichtig zu verstehen, wie verschiedene Zeitdrucksituationen und Kartentypen die Qualität von kartenbasierten Entscheidungen beeinflussen, und welche Rolle die Eigenschaften der Entscheiderinnen und Entscheider in dieser Hinsicht spielen.

Viele empirische Kartenstudien haben die Effektivität und Effizienz von Entscheidungen mit verschiedenen Karten untersucht. Jedoch ist wenig darüber bekannt, inwiefern Zeitdruck in diesem Zusammenhang relevant ist. Die vorliegende Arbeit versucht, diese Forschungslücke anhand von vier kontrollierten Benutzerstudien zur kartenbasierten Entscheidungsfindung unter Zeitdruck zu schliessen. Die Erkenntnisse dieser Experimente werden durch verschiedene Experteninterviews mit Personen ergänzt, die in ihrem täglichen Leben kartenbasierte Entscheidungen unter Zeitdruck treffen.

In diesen vier Benutzerstudien wird der Effekt von Zeitdruck als unabhängige, exogene Variable gemessen. Experiment I widmet sich dem Einfluss von Zeitdruck auf die Präferenzen für verschiedene Kartentypen und Interaktivitätswerkzeuge für die Routenwahl. Die folgenden Experimente untersuchen, inwieweit Zeitdruck die Genauigkeit von sowie das Vertrauen in Entscheidungen bei der Routenwahl mit Strassenkarten und Satellitenbildern (Experiment II) und der Hangneigungsermittlung mit 2D- und 3D-Karten (Experiment III) beeinflusst. Schliesslich wird der Effekt von Zeitdruck auf die Benutzerinteraktion und auf verschiedene Entscheidungen mit interaktiven Karten erforscht (Experiment IV).

Die Ergebnisse dieser Arbeit deuten darauf hin, dass sich Zeitdruck auf kartenbasierte Entscheidungen in vielerlei Hinsicht auswirkt. Bei der Routenwahl beeinflussen verschiedene Zeitdrucksituationen das Vertrauen in die Entscheidungen mehr als die Genauigkeit. Im Experiment zur Hangneigung nehmen Genauigkeit von und Vertrauen in Antworten signifikant ab, wenn der Zeitdruck auf die Testpersonen abnimmt. Weiterhin wirkt sich Zeitdruck auf die Häufigkeit der Benutzerinteraktionen sowie auf die Präferenzen für Kartentypen und Interaktivitätswerkzeuge aus.

Darüber hinaus legen die empirischen Ergebnisse dieser Arbeit nahe, dass abstrakte 2D-Karten mit weniger visuellem *clutter* geeigneter für kartenbasierte Entscheidungen unter Zeitdruck sind als realistische 3D-Darstellungen, und dass die Vertrautheit mit Kartentypen in dieser Hinsicht eine wichtige Rolle spielt. Die vorliegende Arbeit leistet einen Beitrag zu einem besseren Verständnis der Frage, wie Menschen Karten in Zeitdrucksituationen einsetzen, und liefert erste Erkenntnisse darüber, was bei der Erstellung von Karten für die effektive und effiziente Entscheidungsfindung in solchen Situationen beachtet werden sollte.

ACKNOWLEDGEMENTS

This thesis would not have been possible without the help and support of many colleagues and friends. It is impossible to adequately acknowledge all the people who contributed to this thesis in different ways. I would like to express my thanks and acknowledgements to some of them in the following paragraphs and inwardly thank all the others.

First of all, I thank my main supervisor Prof. Dr. Sara Irina Fabrikant for giving me the possibility to carry out my research in her group, for always being ready and open to discuss ideas with me, and for her time, trust, support and advice. I am also grateful to my committee members, Dr. Tumasch Reichenbacher and Prof. Dr. Daniel R. Montello, and to my external reviewer, Prof. Dr. Amy Lobben, for numerous helpful comments and suggestions on my work, and to Dr. Amy Griffin for sharing her insights on signal detection theory.

I would like to extend my thoughts to all my colleagues at the GIScience Center at the University of Zurich for providing an excellent, inspiring environment for my research. In particular, I would like to thank the people at the GIVA group, Dr. Arzu Çöltekin, Dr. Christian Kaiser, Dr. Anna-Katharina Lautenschütz, Stefano de Sabbata, Paul Crease, Marco Salvini, Kenan Bektaş, and Sara Maggi, for discussing ideas and giving helpful advice in many different aspects.

I owe many thanks to the numerous participants who were willing to participate in my experiments. In this context, special thanks go to Dr. Christian Häberling and Dr. Olaf Schnabel for giving me the opportunity conduct my experiments in their cartography classes at the Swiss Federal Institute of Technology (ETH) in Zurich.

I would like to express my gratitude to my interview partners at Schutz & Rettung Zürich, Rega, Stadtpolizei Zürich, EWZ, and the Swiss Institute for Snow and Avalanche Research (SLF), for sharing their time and precious insights on map-based decision making under time pressure. In this context, special thanks go to Prof. Dr. Ross Purves (University of Zurich), Rolf Würth, Dirk Sussenberger and Raphael Näf (Intergraph), and Jörg Sparenborg (EWZ), who have all been very supportive in identifying and establishing connections to these interview partners.

Last but not least, I am also indebted to my parents for their constant encouragement, patience, moral support and understanding during my doctoral studies, and all the friends who have shared many ups and downs with me in the previous four years. Thank you for listening to me, being the necessary touch to reality, and sharing life with me besides academics.

CONTENTS

List of Figures

LIST OF TABLES

NOTES

All figures and tables by the author unless otherwise stated.

All internet sources last accessed March 28, 2012.

„... An unavoidable fact about human decision making is that decisions take time, and the amount of time spent making a decision influences the final choice“ (Busemeyer and Townsend, 1993).

1. Introduction

1.1 Problem statement and motivation

In many situations in life, we have to make decisions under time pressure, and the amount of time we spend thinking about a decision often influences our final choices. Often, our everyday decisions do not only have a temporal component, but are also spatially relevant and are made based on maps (see Figure 1).

Examples include activities in life-threatening situations (such as search and rescue, evacuation, firefighting, military decision making operations), leisure activities (such as orienteering or ski touring), or handling complex traffic situations. For example, commuters may have to decide quickly which alternative route to take through congested traffic, hikers might need to rapidly select a different trail using a topographic map, when the weather suddenly deteriorates in the mountains, or a sailor might have to quickly consult a nautical chart when navigating in an area with sudden wind and water level changes.

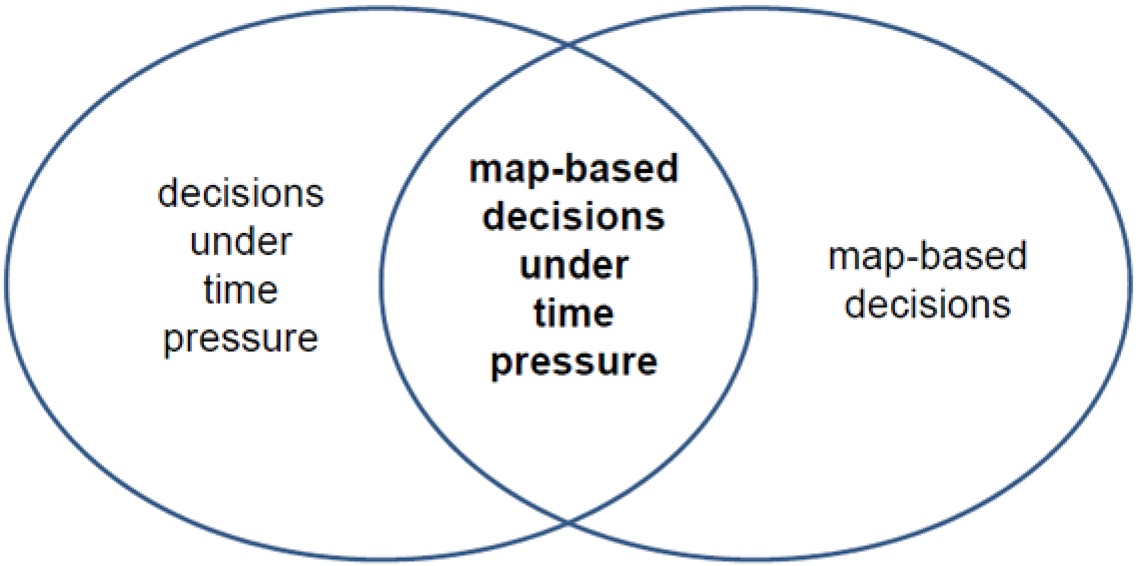


Figure 1: Map-based decisions under time pressure.

Time available for such kinds of map-based decisions can vary enormously. Sometimes a person may have merely a few seconds in which to make a life and death decisions. With both increasing human mobility and increased availability of mobile map devices, it seems crucial to investigate how decision time constraints and display types might affect the quality of map-based decision making under rapidly changing conditions.

Different kinds of spatial displays are employed for decision making: satellite images, fully-interactive globe viewers, sketch maps, road atlases, mobile maps on small-sized PDAs or smartphones, as well as large-scale topographic paper maps. In times of volunteered geographic information, generating different representations of spatial phenomena has become very easy and popular among many, often non-expert users. This development makes it even more necessary to explore which elements make a map most efficient and effective for spatio-temporal decision making under varying usage scenarios.

One particularly common example for map-based decision making under time pressure is road selection. People often do not have much time to think about which route to take on a trip. Imagine driving a car and having to spontaneously select an alternative route due to a traffic jam, and your only decision support consists of a paper map: In this situation, there is no time for elaborate *thinking* about which route is the best one, and most likely you will be under a certain time pressure when having to make this map-based spatio-temporal decision.

This potential effect of time pressure on map-based road selection is particularly important in emergency situations, in which decisions under high time pressure can decide between life and death:

> *„In some circumstances, such as military and emergency response operations, imagery may be the best or only source of spatial information. For such time-critical situations, it is important to know how well an aerial photograph or satellite image can be used successfully as a map.“ (Dillemuth, 2005).*

This quotation suggests that it is important to have a visual display at one's disposal which supports the process of effective and efficient decision making, specifically in time-critical decision making scenarios. In this particular case, the efficiency of a map-based decision is arguably as relevant as its effectiveness. In usability research, effectiveness generally stands for the extent to which a goal is reached, while efficiency measures also consider the effort to reach a goal (International Standards Organisation, 1998). Transferring this concepts to map-based decision making, a map's **effectiveness** is determined by how much it helps map users give accurate answers (irrespective of the decision time spent) with a high confidence, while measures for the **efficiency** of a map-based decision also have to make allowance for the decision time, and report whether map users can give both accurate *and* fast answers with a high confidence. In other words, an efficient map should also be effective, while an effective map does not necessarily have to be efficient. Efficiency might not only involve decision time, but also the perceived cognitive effort or the subjective judgment of how easy it is to make the map-based decision.

The main motivation for this thesis is that time pressure is an under-researched, but relevant independent factor influencing the efficiency and effectiveness of map-based decisions. Various empirical map use studies have looked at the efficiency and effectiveness of decision making with visuo-spatial displays, and focused on the influence of individual or group differences (see related work in Chapter 2), map display types, map designs and interactivity levels. However, very little is known about role the usage condition of time pressure plays in this context. The sheer importance of an efficient and effective map for decision making identifies the need for more generalizable empirical research focused on map use under time pressure.

1.2 Factors influencing the effectiveness and efficiency of maps

Many factors influence the effectiveness or efficiency of map-based decisions (see Figure 2). These factors can be classified and structured employing established cartographic design approaches such as the "pillars of map design" (Fabrikant and Goldsberry, 2005, p. 2) or the "controls of the map design" (Robinson et al., 1995, p. 330). Using these systematic approaches, certain context-related, map-related, and user-related influencing factors can be identified. Context-related factors involve the conditions of use, such as varying map use purposes and decision-making tasks (road selection, slope detection, ski touring) or map use environments (bad weather, time constraints, impaired visibility). Inherently map-related factors take into account the visual appearance, graphical layout and technical limitations of the map itself, while user-related factors consider who is actually making decisions with the map.

This thesis investigates map-based decision making. The specific focus is on four factors, which require special attention (see Figure 2): The context-related factor of decision time, the map-related factors of spatial display (such as varying display types and map designs) and interactivity (different degrees or types of interaction tools), and finally user-related characteristics, among which one can distinguish between individual and group differences. These factors will be further discussed in the remainder of this chapter.

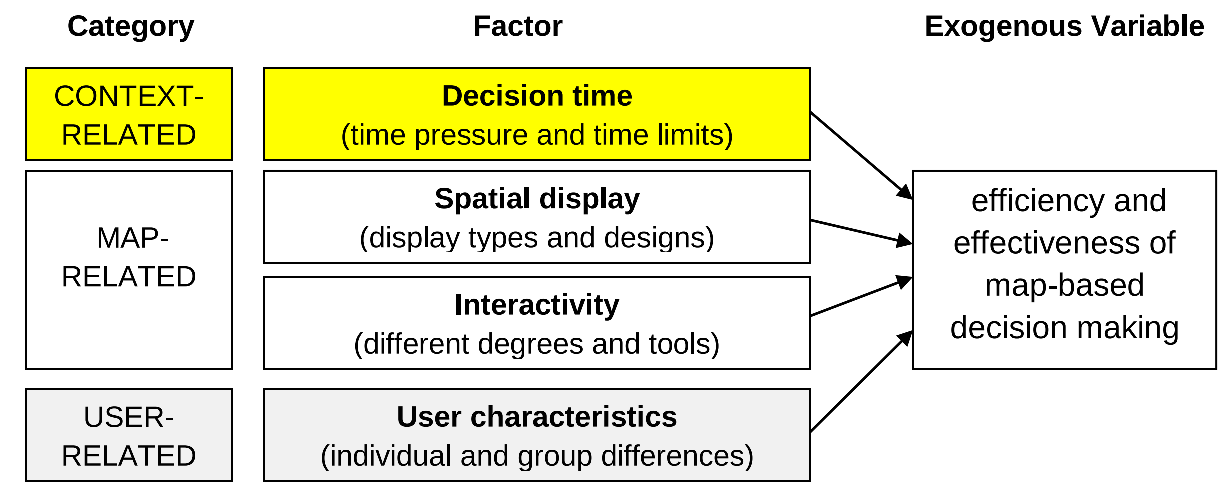


Figure 2: Factors influencing the efficiency and effectiveness of map-based decision making.

To begin, issues of varying time limits and time pressure levels can be summarized under the variable of **decision time.** The concepts of *time pressure* and *time limits* cannot simply be equated, nor can they be disassociated (see also section 2.2). In simplest terms, time pressure is merely a time constraint or, in other words, a temporal deadline by which a certain decision has to be made (Edland and Svenson, 1993). However, this definition neglects different perceptions of time constraints. Some people might regard a certain time limit as pressure, while others might regard the identical time limit as no pressure at all. These competing concepts will be further discussed in the *Related work* (Chapter 2) and *Methodology* (Chapter 3) sections, where time pressure will be operationalized for the context of this thesis.

Map display types and map design issues, such as the information contained on a map and how it is visually represented, are mainly driven by the cartographer's choices. These choices should be influenced by the purpose of the map, the user and technical constraints (Robinson et al., 1995). In this thesis, the variation of **spatial display types** will involve displaying spatial data differently on identically sized displays. These map display types are understood as categories of static depictions of spatial datasets. Examples of such spatial display types are *road maps, satellite images* and *terrain maps* (see Figure 3). Other map display types include *topographical maps* (plane survey sheets), *3D representations*, and *oblique perspective satellite images*. Variations of map display *sizes*, such as large screens vs. small PDA displays, are not within the scope of this thesis, but are certainly relevant for a mobile society. Animated maps are not investigated either in this research.

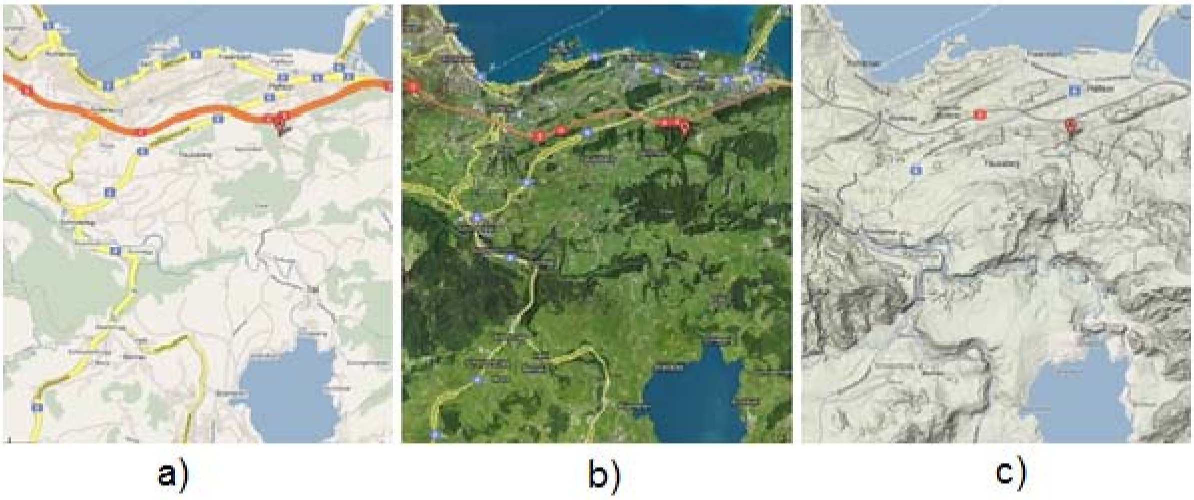


Figure 3: Three different map display types, which are representing the same geographical extent: a) road map, b) satellite image, c) terrain map.

Issues of **map design** take into account that maps of a certain spatial display type may be depicted differently, and thus can vary in their visual charateristics. For instance, a map of the display type *road map* can contain distance information, street names, or different colors for different road types. Moreover, one road map can be visually more complex than another road map, or overloaded with information which is not relevant for the task at hand, or the relevant information might not be saliently displayed. Satellite image maps might differ in their spatial resolution. In other words, map design issues can be regarded as differences "within one spatial display type".

Maps can also vary in their levels of **interactivity.** When assessing interactivity for the task at hand, one has to take into account which type of interactivity is needed, and how it should be implemented (Harrower and Sheesley, 2005). Using Robinson et al.'s (1995) terminology of "controls of map design", interactivity is not only influenced by the purpose of the map, but also by technical limits. In this research, the analysis of interactivity will focus on the four most commonly encountered interaction tools in standard interactive computer map applications, which are *zooming, panning, rotating* and *tilting*, as provided for example by any standard 3D Geographic Information System (GIS), such as *Google Earth*[1] (Ortner, 2011).

Lastly, user-related factors generally take into account *who* is making the decision with a map. When considering **user characteristics,** one can differentiate for example between *individual* and *group differences***.** This distinction is based on the assumption that every individual is in certain respects "like all other men, like some other men, and like no other man" (Kluckhon and Murray, 1953, p. 35). While studies of individual differences often distinguish individuals

[1] http://www.google.com/earth/index.html

by parameters that can be distretely measured (i.e., "like no other man"), such as scores of various abilities (such as spatial or mathematical), studies of group differences emphasize the aspect of individuals being "like some other men" and thus focus on homogenous groups or classes of a common characteristic for their analyses, e.g., by age group or gender (see section 2.6 for a more detailed discussion).

Of course, this is a non-exhaustive list of factors influencing map-based decision making. In *Related work* (Chapter 2) and *Methodology* (Chapter 3), the factors mentioned in this section will be further discussed and operationalized.

1.3 Research questions

The empirical work of this thesis is guided by the following main research question:

What is the effect of time pressure on map-based decision making considering varying spatial display types and map designs, interaction tools, and users with varying backgrounds?

Based on this main research question, several, more specific research questions can be formulated and grouped according to the categories of influencing factors mentioned previously:

Context-related factors (Decision time)

- How are map-based decisions affected by time pressure?

Map-related factors (Spatial display factors and interactivity)

- How do people's preferences for spatial display types and map designs depend on the decision time for map-based decision making?
- How do the efficiency and effectiveness of people's spatial inference and decision making depend on the spatial display type, the map design and interactivity?

User-related factors (User characteristics)

- How do map use preferences depend on individual (e.g., spatial abilities) and group differences (e.g., sex)?
- How does map use performance (i.e., accuracy and confidence) depend on individual and group differences?

I approach these questions within an empirical research framework based on human-subject experiments and expert interviews, which is presented in chapters 3 to 8.

1.4 Relevance for geography and geovisualization

This research is positioned within the field of geovisualization. Geovisualization, which draws upon different approaches from cartography, GIScience, image analysis, computer graphics and cognition (Dykes et al., 2005), is concerned with providing theory, methods and tools for the visual exploration, analysis, synthesis and presentation of geographic data (MacEachren and Kraak, 2001).

The fields of geovisualization and also cartography have always been challenged to produce maps that are as efficient and effective as possible for various usage scenarios and user groups. This challenge implies providing content and visualization methods for maps which can be easily understood, in order to support effective and efficient spatio-temporal decision making. In order to construct cognitively adequate maps for map-based decisions, one has to understand the way people perceive maps; that is, the cognitive processes which determine the relative success or failure of spatial displays have to be well understood.

As a consequence, a special focus in geovisualization is the aspect of cognitive and usability issues, as mentioned in several research agendas for geovisualization (Dykes et al., 2005; Gahegan, 2008; MacEachren and Kraak, 2001; Montello and Freundschuh, 2005; Slocum et al., 2001). Cognitive and usability studies focus on cartographic human factors and cartographic communication (Montello, 2002). These studies are based on the assumption that *"geovisualization is not just about technical issues"* (Dykes et al., 2005, p. 15), and try to explore whether the products from geovisualization *"actually work and under which circumstances this is the case"* (Dykes et al., 2005, p. 6). Within this field of cognitive and usability studies about map design, this thesis particularly focuses on the aspect of *"evaluating the effectiveness of geovisualization methods"* proposed by Slocum and colleagues (Slocum et al 2001, pp. 12 ff).

1.5 Aim and approach of the thesis

The overall aim of this thesis is to bridge the existing research gap between time pressure research and empirical map design and map use studies, and thereby to make a contribution to a better understanding of map-based decision making *under* time pressure.

I propose to fill this research gap with a set of controlled user experiments about map-based decision making under time pressure. In these experiments, I will assess how the context-related, map-related, and user-related factors discussed in the previous section (i.e., decision time, spatial display factors, interactivity, and user characteristics) influence the efficiency and effectiveness of map-based decision making. A special focus in these experiments will be on time pressure, as it is under-researched in the domains of cartography and geovisualization, as discussed in Chapter 2. These experiments will be complemented by expert interviews with professionals in the field of map-based decision making under time pressure.

1.6 Structure of the thesis

The remainder of this thesis is structured as follows: **Chapter 2** provides an overview about the current state of the research on map-based decision making under time pressure. It provides the theoretical basis for the empirical work presented in the remainder of the thesis, which is structured as follows:

In **Chapter 3,** I introduce the methods that will be employed for addressing the research questions. Chapters 4 to 8 contain the descriptions and results of the empirical work. **Chapter 4** focuses on Experiment I, in which I assessed map use (i.e., spatial display and interactivity) preferences for a road selection task under time pressure. **Chapter 5** is devoted to Experiment II, a follow-up experiment about map use performance (i.e., accuracy and confidence) for a road selection task with static spatial displays in a flat urban environment. **Chapter 6** summarizes the results of several expert interviews in the field of map-based decision making under time pressure. In **Chapter 7,** I present a slope detection experiment (Experiment III) with static 2D- and 3D-looking maps, based on the findings of the expert interviews in Chapter 6. **Chapter 8** reports the results of the last Experiment IV in this research, in which different map use tasks were carried out using an interactive virtual globe.

Chapter 9 discusses the main results of this thesis, organized by the different factors influencing the effectiveness and efficiency of map-based decisions, as introduced in Chapter 1. **Chapter 10** concludes the thesis with a summary of the main findings of this research and an outlook on future research directions in the field of map-based decision making under time pressure.

2. Related work: Theory and state of the research

Related work in map-based decision making under time pressure can mainly be divided into two strands: General decision making approaches on the one hand, and map use studies about the effectiveness and efficiency of decision making with visuo-spatial displays on the other hand. Only little research has been conducted so far which has tried to combine these two strands.

This chapter reviews and discusses the current state of the research in this field, in order to identify research gaps and to build a solid theoretical basis for the experimental portion of this thesis. This literature review is structured according to the factors that influence the efficiency and effectiveness of map-based decision making, as presented in section 1.2 and shown in Figure 2.

Following this structure, I will first discuss the **context-related** factor of time pressure and its role in the broader context of decision making studies, mainly outside the fields of GIS and cartography. Afterwards, I will present how decision time has been used in studies in GIS and cartography, mainly in the field of road selection, and as a performance measure.

In a next step, I will present the influence of **map-related** factors by highlighting numerous studies in which the efficiency and effectiveness of decision making with different visuo-spatial displays and degrees of interactivity have been investigated. Thereby, I will identify research gaps about the effects of different spatial display types, map designs and interactivity on map-based decision making, which should be further investigated under time pressure.

Finally, I will survey the literature about **user-related** factors, that is, to what extent individual and group differences are relevant for decision making with spatial displays.

2.1 Context-related factors I: Time pressure outside cartography and GIS

2.1.1 Limitations to optimal decision making

Time pressure is one of many factors that contribute to the phenomenon that humans do not always make fully rational, optimal, or perfectly accurate decisions in reality (Payne, 1982). Hence, the broader context of human decision making studies seems a good starting point before focusing specifically on time pressure.

Empirical studies of human decision making have had a long history in many disciplines, such as mathematics, statistics, micro-economics, and psychology (Kahneman and Tversky, 2000). While early decision making studies have assumed that humans make fully rational decisions

(Bentham, 1789; Pareto, 1906), and rationality is still regarded as the most common portrayal of decision making (March, 1994), many authors have refined this heuristic model, which equates the human decision-maker with the utility-maximizing "*homo oeconomicus*", (e.g., Tversky and Kahneman (1981)).

One of the main critiques of fully rational or optimal decision making is that humans only rarely have full and relevant information on which to base their decision and can often not be certain about the consequences of their actions. Hence, it seems more appropriate to describe human-decision making using the model of limited or *bounded rationality.* This concept, coined by Simon (1959), tries to characterize human decision making in economics in a more realistic way than previous, mathematical approaches did. Bounded rationality implies that decision making is impaired by several factors. These factors include the limitations of knowledge of alternatives and consequences, constraints in attention, memory, comprehension, communication, and also the infinite amount of time for human decisions. As Gigerenzer (2002) puts it, human decision-makers cannot be described as *"Laplacean superintelligence equipped with unlimited resources of time, information and computational mind"* (p. 37), as it had often been assumed in early studies in economics, cognitive science, and biology. Moreover, humans tend to look rather for "good enough" than the "best possible" solution for solving a problem (March, 1994).

A second critique addresses the disregard of potential individual or group differences in standard economic rationality theories. That is, in reality not all human decision-makers will react to the same conditions in the same way, as for example "expected value theory" suggests (Yates, 1990). For instance, people vary in their individual expected utility functions due to personal preferences (Friedman and Savage, 1948), in individual decision weights assigned to certain probabilities (Tversky and Kahneman, 2000), or in their behavior under risk, which can be described as risk-averse or risk-seeking (Rabin, 2000). Such individual differences in decision making are also reflected in decision making behavior under time pressure, as shall be seen in the next section.

2.1.2 Definitions: Time pressure vs. time constraints

Introducing the concept of time pressure, it should first be noted that identical time limits might cause different perceptions of time pressure and stress among different people. In other words, while some people might perceive a certain time limit as time pressure, other people will not feel any time pressure by this identical temporal deadline. A clear distinction between the concepts of *time constraints* and *time pressure* has been suggested by Ordoñez and Benson III (1997): A *"time constraint exists whenever there is a time deadline, even if the person is able to* *complete the task in less time"*, while *"time pressure indicates that the time constraint induced some feeling of stress and created a need to cope with the limited time"* (p. 122). As a consequence, it is possible to be under a time constraint, but not under time pressure. For this reason, Johnson et al. (1993) suggest the number of Elementary Information Processes (EIP) as an alternative indicator for time pressure. These authors also argue that time pressure is indeed perceived differently by individuals, and thus will often vary among people and participants of controlled experiments, who are exposed to identical time limits. However, it seems cumbersome to transfer the concept of EIP to map-based decision making. Firstly, EIP is difficult to identify for a given task with a given map in a given situation. Secondly, map-based decisions usually can be made in multiple ways, which would involve different processes and thus require different amounts of time and effort.

2.1.3 Negative effects of time pressure on decision making (Speed-accuracy tradeoff)

The effect of time pressure on decision making has been researched to a great extent by cognitive, developmental, and personality psychologists, as well as by human resource researchers and economists (see Förster et al. (2003) for an extensive review). Although a single tradition of research and an overarching theory seem to be missing (Edland and Svenson, 1993; Maule and Svenson, 1993), it is widely accepted that decision time influences the quality of decisions (Svenson et al., 1990), and that the negative effect of time pressure on decision making is robust and consistent (Ahituv et al., 1998; Pew, 1969).

Several authors (Maule and Edland, 1997; Wickelgren, 1977) have called this phenomenon characterizing the negative of effect of time pressure on decision accuracy **speed-accuracy trade-off**, suggesting that time pressure reduces the overall quality of a decision, and that people give more accurate answers when they have more answering time. Still, the exact characteristics of the negative effect of time pressure on decision accuracy actually seem to be quite contested.

Figure 4 shows speed-accuracy relationships for different types of tasks. The phenomenon of a trade-off is clearly visible in this figure, as in each of the six different experiments presented by Johnson et al. (1993) the accuracy of answers decreased with increasing time pressure. However, the different slopes and intercepts of the curves also suggest that the shape of the speed-accuracy trade-off varies between the tasks. The steeper the slope in this diagram, the stronger relative accuracy is affected by time pressure, and therefore the more striking the speed-accuracy trade-off is for this experiment. Johnson et al. (1993) regard task complexity as the main driver of the speed-accuracy trade-off: That is, the more complex a task, the more likely the occurrence of the speed-accuracy trade-off.

Hwang (1994) states this the other way round: He argues that for detecting any effect of time pressure on response accuracy, time pressure must be strong enough to change the difficulty of the task. In other words, only if participant performance decreases at a certain time limit, can one say the task is getting more difficult under this time limit.

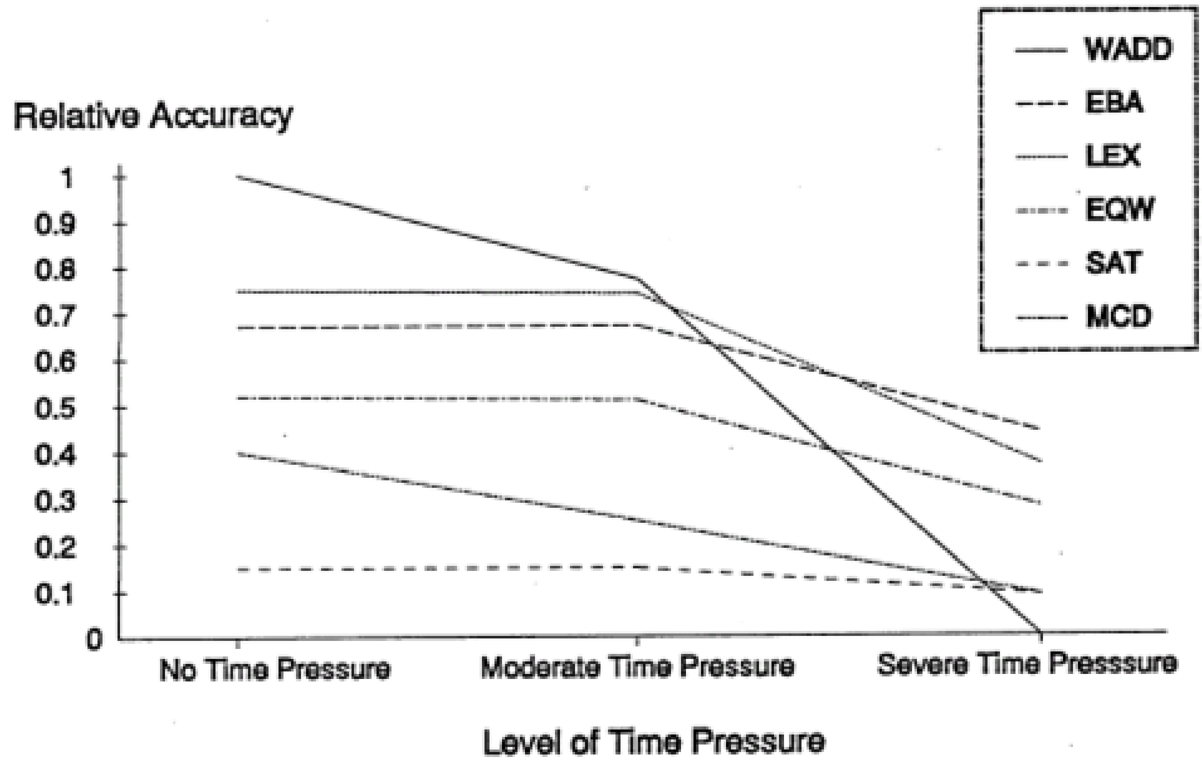


Figure 4: Different manifestations of the speed-accuracy trade-off for different experiments (Johnson et al., 1993).

A simple calculation example can be used to illustrate this hypothesis: Let us think of an experiment that consists of several mental arithmetic multiplying tasks with integers between 10 and 100, under a time constraint of 10 seconds for each multiplication. If we now extend the time constraint for the calculations to 60 seconds per task, one can assume that overall response accuracy will increase, as most people can solve this task perfectly accurately within one minute. However, if the response time limit is increased from 60 seconds to ten or even twenty minutes, one would not expect any further increase in response accuracy, because one would not assume that these tasks are getting easier with more than 60 seconds response time. Therefore, if one would use the three response time limits 60, 120 and 180 seconds for this multiplication experiment, one would probably not detect any speed-accuracy trade-off. This example indicates that there are certain caveats and implications when setting response time limits and talking about effects of time pressure on decision making: Only if the response time limits are set in a way that task complexity or difficulty is increasing, can any speed-accuracy tradeoff be found.

2.1.4 Positive effects of time pressure on decision making (Inverted U-shaped curve)

Several researchers have also reported positive effects of time pressure on decision making. For instance, Andrews and Farris (1972) studied the performance of NASA scientists and engineers, and found that their performance increased with time pressure. Performance only decreased when time pressure became too severe. This result of a performance increase with time pressure was replicated in a similar study by Peters and O'Connor (1980) involving commercial bankers as decision makers. Hwang (1994) argues that perhaps the best way to describe the interaction between decision performance and time pressure is not a linear relationship, but an **inverted U-shaped curve**. As he states it, *"increasing time pressure leads to better performance up to a certain point, beyond that point more time pressure reduces, rather than increases, performance."* (p. 198).

Therefore, it can be summarized that there are two contesting hypotheses about the effect of time pressure on the accuracy of decision making: The first "speed-accuracy trade-off" hypothesis suggests that accuracy is generally decreasing with more time pressure, which implies that accuracy is highest without any time pressure. The second "U-shaped curve" hypothesis suggests that accuracy reaches a maximum somewhere between *no time pressure* and *severe time pressure*. With my empirical work, I empirically investigate which of these two hypotheses can be maintained for *map-based* decisions, guided by the research question: Can the patterns of the speed-accuracy trade-off and the inverted U-shaped curve also be found in map-based decision making under time pressure?

2.1.5 Speed-confidence trade-off

Time pressure does not only have an effect on the accuracy of decisions, but also on the confidence in decisions. The existence of a **speed-confidence trade-off** has been demonstrated by several authors studying investment analysis and typical situations for managers (Maule, 1998; Maule and Andrade, 1997; Smith et al., 1982). In analogy to the speed-accuracy trade off discussed in the previous section, the speed-confidence trade-off implies that the confidence of decision-makers in their decisions decreases with increasing time pressure. Garrett (1922) already considered speed-accuracy and speed-confidence trade-offs in the early 1920s. He assumed that a similar effect on confidence would be a plausible consequence of a speed-accuracy trade-off. Pew (1969) further developed this idea and argued that response time should be correlated with relative confidence on a logarithmic scale.

However, Zakay (1985) reports that time pressure does not change participants' confidence for a choice process, but rather influences their strategies in favor of the use of simpler rules. These simpler rules and strategies people use in order to cope with time pressure include *acceleration* (an increase of the speed of processing information), *filtration* (high priority is given to the most important information), and *omission* (less relevant aspects of the decision making process are left out) (Miller, 1960). Transferring this concept to map reading, it would imply that, when making decisions with a map under time pressure, only the most relevant information of a map would be perceived, while less relevant map items would be left out.

2.2 Context-related factors II: Time pressure in map-use studies

Having introduced the topic of decision making under time pressure, in the following section I will establish the link between time pressure outside the field of GIScience and the work that has focused on decision making with maps. First, I will present road selection as a typical map-based decision making task under time pressure. Then, I will discuss how response time has been employed in empirical map use studies.

2.2.1 Route selection and wayfinding

As mentioned in the introduction, one very typical task for map-based decision making under time pressure is route selection during navigation. Important issues in this context include different road selection strategies people pursue and cognitive biases they might have solving this task (Golledge, 1995). The fact that humans are not optimal decision-makers (as discussed in section 2.1.1) is also reflected in route selection tasks (Pingel, 2010). In a study using navigation devices, Pingel and Clarke (2005) make the case that people are only "boundedly rational" and make suboptimal decisions when selecting routes, as they only use navigation systems as is minimally necessary, and not to the extent that would optimize their route selection choices.

Examples of route selection strategies include (but are not limited to) asymmetric route choice preferences and minimizing mental effort (Christenfeld, 1995), road climbing (i.e., a preference for long and straight routes, see (Bailenson et al., 1998), or the preference for simplest instead of the shortest paths (Duckham and Kulik, 2003). Hochmair and Rinner (2005) distinguish between non-compensatory and compensatory decision rules, which has been developed in time-pressure research outside of geography (Maule and Andrade, 1997). These authors argue that users will prefer non-compensatory rules when under time pressure, that is, rules where a poor evaluation on one attribute cannot be compensated by a positive evaluation on another attribute (Edland and Svenson, 1993).

Several authors have shown the influence of the design of spatial displays on route choices. For instance, Gill (1993) has observed that participants make different route selection choices with different map designs. His results suggest that people prefer and perform better using maps in which the road classes were depicted by distinct and unambiguous line ordering and colors. Hochmair (2009) has demonstrated the influence of map design on route choices made with public transportation maps. In his study, he could show that the additional information on "headway maps" allows for planning faster routes than when participants used a schematic map without map annotation.

Brunyé and colleagues have also shown cognitive biases in route selection: They reported that participants perceive travelling north as travelling uphill and therefore significantly prefer southern over northern routes, while the preference differences between western and eastern routes was minimal (Brunyé et al., 2010). At this point it is an open question whether time available for a map-based decision might influence the preference for a certain route selection strategy, and how it might enforce cognitive biases.

2.2.2 Response time in map-use studies

In empirical cartographic research, response time has typically been employed as a dependent variable (i.e., efficiency measure) to evaluate cartographic design principles (Garlandini and Fabrikant, 2009; Ishikawa and Yamazaki, 2009; Lloyd and Bunch, 2003) rather than as an independent variable.

For instance, Dillemuth (2009) used response time under time pressure as an additional success measure besides accuracy and confidence in an empirical work on navigation with small displays. She provided participants with the identical time limit of five minutes, but did not vary time pressure as independent variable. Her results indicate that response time of participants is generally decreasing with a larger map extent, because participants have to interact more with the map when the map extent is small. In particular, her participants had to perform more panning operations with a small map. Kuoa et al. (2006) measured response time with different display types and visual tasks (Parallel Coordinate Plots, SOM component plane displays and 2D/3D projections), and showed it was very task-dependent as for which representation participants required more response time.

Studies which have varied time pressure as a controlled, independent variable or factor, in order to assess its effect on map-based decision making, seem to be very sparse in the geovisualization literature. Baus et al. (2002) suggest considering time pressure when designing displays of mobile devices for pedestrian navigation. They argue that different travelling speeds during navigation create varying time pressure situations, which in turn should lead to different user requirements for navigation displays. They contend that different content should be displayed on a map used in different time pressure conditions. In another study involving user motivation in navigation, Srinivas and Hirtle (2010) offered a reward to one participant group as an incentive for faster task completion, while the other "control" group was not given any incentive to reduce task completion time. Their results show that "more motivated" participants completed the routes significantly faster than the participants in the "control" group. This might be another indication of the positive effect of time pressure on map-based decisions, as described in section 2.1.4.

2.3 Map-related factors I: Spatial display types and map design

Making a decision with a map is a process which consists of different steps. Some of these steps are typical of any map use situation: seeing (perceiving) visual information, then processing it, and finally interpreting, understanding and attaching meaning to it (MacEachren, 1995). The more complex the map use task, the more information has to be perceived, processed and interpreted, and thus the more time-consuming are these steps (Board, 1978; Brophy, 1980). After having gathered the information from the map, the map-based decision maker can actually proceed to the steps which are inherent in any decision-making process. These steps might consist of evaluating different solutions and finally choosing the best option (Yates, 1990). The decision-making steps also require a certain amount of time, which is influenced by the complexity of the decision-making task (Hwang, 1994), as shown in section 2.1. If we assume that it is better to have more time available for the decision-making process, it is particularly important to identify map-related factors that enhance the efficient (i.e., accurate and fast) visual processing of the map (Slocum et al., 2001).

The influence of map-related factors on the effectiveness and efficiency of decision making with maps has been discussed in several theoretical and empirical studies in geovisualization, cartography, and psychology. The main issue in this context is how differences in levels of realism and abstraction, dimensionality of displays, or the cognitive adequacy of the representation, influence the accuracy and confidence of human decision making.

First arguments for evaluating how people look at and interpret maps date back to as early as the turn of the last century when Eckert (1921) made a claim for the application of psychological research in cartography in order to cope with the subjectivity involved in map communication. Since then, numerous psychophysical and cognitive studies about the efficiency and effectiveness of maps have been conducted (see Montello (2002) for a comprehensive overview).

In the following, I will survey related studies that have investigated the effect of map-related factors on map-based decision making.

2.3.1 The effect of map-related factors on effective and efficient decision making

One key issue regarding the influence of map-related factors on the efficiency and effectiveness of map-based decision making is the extent to which different degrees of realism and abstraction influence response accuracy (i.e., effective decision making) or response time (i.e., efficient decision making). Numerous prior empirical studies in cartography have investigated how different map designs might influence human visuo-spatial inference and decision making, typically depending on a specific map use task (Fabrikant and Lobben, 2009), and often draw upon visual search theories (Dobson, 1983; Lloyd, 1997), or cognitive load theory (Bunch and Lloyd, 2006). Overall, there are very mixed results regarding the relative advantages or disadvantages of realistic or more abstract representations, or 2D vs. 3D displays. Differences in map use with different spatial display types generally seem to depend on the task:

For example, Dillemuth (2005) compared the performance of participants with two different display types for small mobile displays: firstly, an abstract generalized map, and secondly, a realistic satellite image. Results from this research, which has been carried out in the field with mobile devices, demonstrate that the generalized map was more successful than the satellite image, in terms of shorter time to route completion and less browsing. One of the explanations for this is that the higher level of detail prompted subjects to zoom closer into the map. The degree of success however depended on participants' spatial abilities (see section 2.5), as participants with high spatial abilities also performed significantly better with the generalized map, while there were no such strong correlations between spatial ability and task performance with the satellite image. In terms of accuracy, there was no difference between the two display types. These results indicate that, while both map types might be equally effective, there is a tendency that decision making might be more efficient with the abstract map.

Studying aviator navigation performance, Smallman et al. (2001) have shown that users' search time for selecting aircraft which meet certain criteria was significantly faster with 2D maps than with 3D-looking spatial displays. In related work on the design of cockpit displays, Thomas and Wickens (2006) found no significant performance differences in participants' accuracy and response times between 2D co-planar and 3D perspective displays. Kirschenbauer (2005) has conducted a user study comparing topographic maps and an auto-stereoscopic 3D display. She demonstrated that the true 3D map can have certain benefits, such as an increase in perception and an easier understanding of spatial phenomena.

Coors et al. (2005) evaluated small-screen 3D and 2D mobile navigation aids, and found that the majority of the participants had a positive attitude towards 3D. Their participants mentioned that 3D maps were generally a good idea, but also that 2D was already sufficient for mobile navigation. However, participants' response times were significantly slower with 3D maps compared to 2D maps. This suggests that 3D displays in the context of navigation might be more suitable when having more time available for decision making, but less useful under time pressure.

Many studies have compared the usefulness of more abstract contour maps and more realistic 3D models or hillshading maps. Philips et al. (1975) have found that it is very task-dependent which map performs better. They showed that there was practically no performance difference between contour maps regarding relative and absolute heights. However, in tasks that involved visualizing the landscape, hillshading maps outperformed contour maps and thus seem to enhance efficient decision making. Potash et al. (1978) compared contour maps, shaded relief maps and layer tint maps for several tasks, including slope identification. These authors found that, aggregating all tasks, participants performed worst with shaded relief maps, which suggests that adding shaded relief to contour line maps does not lead to faster or more accurate responses.

2.3.2 Spatial display types, preferences and confidence

Another issue regarding map-related factors is the potential discrepancy between user preferences or confidence and actual task performance.

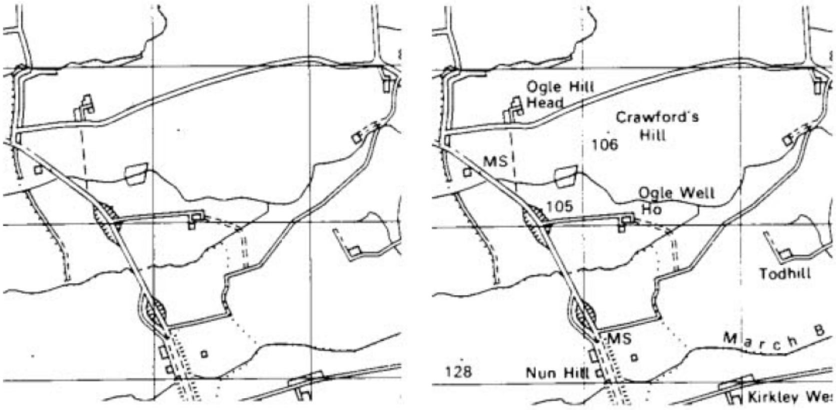
As an example, Canham et al. (2007) and Hegarty et al. (2009) have shown that users tend to prefer more realistic, 3D-looking weather maps that on the surface seem to contain more information for the decision making task at hand than more abstract 2D maps. However, while users prefer 3D, these displays do not necessarily seem to positively influence users' task performance. In fact, Hegarty and colleagues (2009) found that performance was generally better with the less realistic-looking maps, while users' preference ratings indicated just the opposite. These results can be interpreted as *"another good, empirically validated illustration of the common-sense notion that what people think they want is not always what is best for them"* (Fabrikant and Lobben, 2009, p. 141). According to Hegarty and colleagues, "naïve cartographers" seem to prefer 3D displays to 2D displays, and also seem to prefer more realistic depictions to simpler, more abstract ones. Cartographic design theories and principles, however, aim for reducing graphic complexity (Bertin, 1967; Tufte, 1983).

This phenomenon of a relatively high confidence in realistic displays has also been coined "naïve realism" by Smallman and St. John (2005). As a consequence of this naïve realism, their participants preferred realistic displays, even if it hindered task performance in a task related to aviation. Smallman and St. John found that people seem to overestimate their ability to extract relevant information in general. Hegarty et al. (2010) discovered that novice users tend to favor realistic displays more than experienced users, and Boughman (2005) showed that confidence in road distances was higher in photorealistic maps. This over-confidence in realistic depictions has been replicated by different authors (Fabrikant and Boughman, 2006; Hegarty et al., 2009; Zanola et al., 2009). Nevertheless, there are also examples where users preferred more abstract maps to more realistic ones. For instance, in a study with contour maps, shaded relief maps and schematic maps without elevation information, Soh and Smith-Jackson (2004) have found that users prefer the a map with only contour lines over a more realistic shaded relief map.

2.3.3 Map design: Complexity and clutter

Maps of the same spatial display type can be designed differently. There are several graphical ways in which geographic data can be presented, and if a map is to be effective, the representation must be carefully chosen (Robinson et al., 1995). During every design process, the pillars of map design, that is, map theme, map purpose and map audience or user (Fabrikant and Goldsberry, 2005), should be considered, in order to produce a map which is effective and efficient for the task at hand.

Several concepts have been developed for measuring and evaluating the effectiveness and efficiency of certain map designs. These concepts include variations in how densely information is displayed – e.g., clutter, see (Palmer, 1994; Rosenholtz, 2001; Rosenholtz et al., 2005) – or how salient or cognitively adequate the thematically relevant information is presented (Fabrikant and Goldsberry, 2005; Swienty et al., 2008). Moreover, maps can vary in complexity. Fairbairn (2006) introduces different metrics for measuring the complexity of maps, of which the simplest one is the number of items on a map (see Figure 5).



CRAWF1, 1:50,000   CRAWF3, 1:50,000

Figure 5: Two displays with a different number of items: The right one contains elevation information and labels and can therefore be regarded as more complex than the one on the left (Fairbairn, 2006).

However, the number of items is not always straightforward to measure. As an example, what would count as an item in Figure 6? Every single raindrop, every single state, or rather the whole layer of rain or states? And how are the shaded relief, the contour lines for the air pressure and the clouds measured correctly?

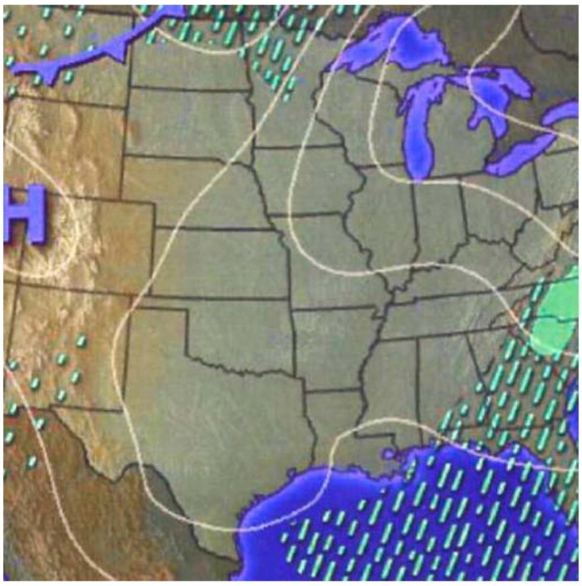


Figure 6: A weather map example suggesting that measuring the number of items on a map is not always straightforward (Rosenholtz et al., 2007).

Due to this problem, more complex, but arguably also more helpful measures for complexity have to be considered. These measures include spatial statistical measures, entropy and image-based indices, which have certain advantages for the analysis of both raster and vector maps. Fairbairn claims that the compression ratio is the most valid measurement of complexity. However, this measure also depends on the task, and its true nature is user-specific.

For this thesis, visual clutter measures, as introduced by Rosenholtz et al. (2007), seem to be more adequate predictors for human map-reading performance than Fairbairn's complexity measures. In the terminology of Rosenholtz et al., clutter relates to the degree of perceptual organization of information in a display. The more organized a display, the less visual clutter (detracting information) it contains for a given task. Rosenholtz and colleagues introduce several metrics of visual clutter, such as *feature congestion* and *subband entropy*, and show that these metrics correlate with search performance in complex imagery. The negative effects of visual clutter on perceptibility have been empirically validated in early research (Dobson, 1980), and Klippel et al. (2006) regard visual clutter even as the "biggest threat to easy perception" (p. 120).

Rosenholtz et al.'s first measure, feature congestion, captures increased color variability, and it can therefore inform about the difference between color, texture and orientation clutter. It is *"based on the analogy that the more cluttered a desk is, the more difficult it would be to add an attention-grabbing note to the desktop"* (Rosenholtz et al., 2007, p. 18). Transferring this concept to this study about maps, this would imply that the more cluttered a map is, the more difficult it would be to add an attention-grabbing piece of information to the map.

Secondly, subband entropy emerges as an even better predictor for the efficiency of map-based decisions than feature congestion. Subband entropy, also empirically validated with spatial displays, is *"based on the notion of clutter as related to the efficiency with which the image can be encoded and inversely related to the amount of redundancy and grouping in the image"* (Rosenholtz et al., 2007, p. 18). The higher the subband entropy measure for a display (i.e., the more clutter), the less computationally efficient the extraction of information encoded in the image (Simon and Larkin, 1987). Thus, subband entropy seems to be a good predictor for human map-reading performance under time pressure.

One should note that certain interaction modes can reduce the complexity and clutter of a visual display. For instance, interactive maps might provide the user with the possibility to "switch off" layers with irrelevant information. In this case, the complexity or clutter of a map can only be calculated for the current "snapshot" of the visualization. Interactive maps will be discussed in the next section.

2.4 Map-related factors II: Interactivity and human-map interaction

In order to produce effective and efficient maps, it is critical to understand how humans interact with maps (MacEachren, 1995). The possibilities to interact with visuo-spatial displays have been growing steadily since the introduction of computers. Thus, it should be investigated how interactivity can enhance effective and efficient map-based decision making.

In the first part of this section, I will review the history and development of virtual globes as a typical example of interactive maps, define the properties of interactive maps, and review relevant empirical studies on human-map interaction. In the second part of this section, I will focus on one particular human-map interaction type and explain why it requires further investigation within the scope of this thesis.

2.4.1 Virtual globes and interactivity

For several thousand years, static planimetric maps have been the state of the art for map-based decision making (Moellering, 2007). However, in the last decades, myriads of computer-based interactive cartographic products have been developed and emerged as an alternative map type for decision making. The first theoretical and practical considerations about virtual globes and interactive 3D cartography (Kraak, 1993; Moellering, 1980) have been made in times where the conceptual and technical hurdles were still so high that the design, let alone the efficient usage, of interactive virtual globes were practically impossible.

Due to technical progress, such as the increase in processing power of computers, and the availability of user-friendly interfaces, the last 20 years have seen a dramatic popularity increase of interactive maps, which have lead cartography into the digital and electronic age (Robinson et al., 1995). In particular, so-called "virtual globes" or "globe viewers" have introduced the concept of fully interactive 3D maps to a wide audience (Riedl, 2006). These globes visualize the Earth *"as a three-dimensional globe that one can fly above"* (Schultz et al., 2008, p. 28). One landmark event in the history of virtual globes took place in 2005, when the free virtual globe Google Earth was launched. Schöning et al. (2008) state that Google Earth has already been downloaded more than 100 million times in the first fifteen months after the initial release.

Virtual globes provide users with the possibility of interacting with the map with implemented interaction tools, such as panning, zooming, rotating and tilting (Schultz et al., 2008). Current state-of-the-art interactive virtual globes have implemented these three-dimensional visualization utilities proposed by Kraak (1993), allowing for *"geometric map transformations such as rotation, scaling, translation and zooming to position the map in 3-d space with respect to the map's purpose and the phenomena to be mapped"* (p. 193). While several design considerations for 3D cartography have been made, and some empirical studies have focused on 3D cartography and visualization (e.g., the works by Moellering, 1980, and Kraak, 1993), the actual effectiveness or efficiency of human-map interactions with virtual globes has not been widely studied by scientists.

At the beginning of this millennium, Cartwright et al. (2001) noted that overall guidelines and theories of interactive geovisualization were missing and that cartographers had only “very little knowledge of the impact of interactivity on how people think or make decisions with interactive environments” (p. 56). Crampton (2002) tried to bridge this research gap by proposing a typology of interactivity. He identified different levels of interactivity, which can be ranked on an ordinal scale, based on the power of interactivity. Using his framework, the virtual globe interaction tools zooming, panning, rotating and tilting belong to the lowest level of interactivity. This level is called “interaction with the data representation”, and does not involve more powerful interactions such as interacting with the temporal dimension or contextualizing interaction. In the terminology of camera motion, zooming in and out can be regarded as translations along the z-axis, while panning represents a horizontal camera motion, and tilting a vertical camera motion (Abend et al., 2011). Harrower and Sheesley (2005) identify zooming and panning as key components of every image browsing tool, including maps, and contend that these two tools have even become ubiquitous for every computer user. This suggests that among the four interaction tools tested in this thesis, zooming and panning can be regarded as more important for map interactivity than rotating and tilting.

The work of Abend et al. (2011) is one of the first studies which focused on how people actually navigate on-screen when using virtual globe viewers. These authors found that, when navigating with virtual globes, users tend to use mixed directions of camera movements, that is, they use a mix of interaction tools rather than just one tool. In addition, their participants had some preferences to retain the north-up orientation of the map, while users with some experience in 3D graphic programs were more likely to tilt the view.

A remaining question is to what extent these novel interactive tools are actually efficient and effective for different map use purposes (Fabrikant, 2005). Research on map interactivity has mainly focused on animated maps so far (Harrower et al., 2000; MacEachren et al., 1998), or on other interaction tools than the ones which are implemented in current 3D globe viewers, such as classification, brushing and focusing (Andrienko et al., 2002).

Furthermore, there is an ongoing debate as to what extent people actually benefit from the possibilities of interacting with a visual display in general. While several studies on visual object recognition (James et al., 2002) or acquiring spatial knowledge in a virtual environment (Peruch et al., 1995) have found significant advantages of providing interactivity to users, other studies on navigating in desktop and virtual environments were not able to detect any benefits of interactivity. There are even examples of studies showing that participants who were searching for structure in 3D data (Marchak and Zulager, 1992) performed worse when provided with possibilities to interact. In a study on user interaction with a 3D visualization for inferring and drawing cross sections, Keehner et al. (2008) have demonstrated that providing participants with active controls of the 3D display did not necessarily enhance their task performance. Having participants passively watch a movie showing the examples of optimal movements by good task performers was already sufficient to improve their task performance to a level that was equal to that of participants directly interacting with the 3D display. Their results indicate that seeing the task-relevant information is more important in this context than allowing people to interact with the display, regardless of whether this information is obtained actively or passively.

While only very little is known on how people use static maps under time pressure and interactive maps in general, even less seems to be known about how people use *interactive maps under time pressure*. Studies on interactive map use under time pressure can contribute to bridging the research gap regarding how humans may gain spatial information from virtual globes. This was identified as an important research issue by Schöning et al. (2008), who have conducted one of the first user studies with a multi-touch virtual globe.

2.4.2 Map rotation

Humans cannot only interact with fully-interactive digital or "virtual" geovisualization environments, but, to a lesser extent, also with static resources (Krygier et al., 1997), such as hardcopy paper maps (Crampton, 2002). The concept of map rotation, which is one of the standard tools implemented in interactive 3D maps, stems from physically rotating paper maps for navigation purposes.

The phenomenon that some people rotate maps so that they are in line with their forward view is called track-alignment. Another common strategy of orienting the map with north at the top is defined as *North-up alignment*. Aretz and Wickens (1992) investigated this at a time when rotating digital maps was not as common as it is now. These authors demonstrate that humans are more accurate in map localization tasks when the map is "track-up". This advantage of track-up displays might have influenced the design of current state-of-the-art car navigation systems, in which the map orientation is in most cases also "track-up" and not "north-up". If a map is not track-up, humans have to perform mental map rotation to bring the map back into congruence with the forward view. This mental rotation can be regarded as requiring additional cognitive load.

Paper map rotation issues have also been researched by Lobben (2004), investigating which map-reading strategies people use when navigating in real-world environments. When travelling with printed paper maps, some people tend to physically rotate maps, instead of "rotating the environment" mentally. However, Lobben (2007) also points out that the relationship between mental rotation ability and the actual map-use strategies is still unclear. It is not well understood whether somebody who is good at mentally rotating objects will or will not rotate a paper map physically. In her study, Lobben found first evidence that people with a poor score at mental rotation (in her case, measured by a map rotation task) tend to rotate the map in a physical environment during navigation. Mental rotation will be covered in more detail in the *User-related factors* section of this chapter (section 2.5.2).

For this study, previous work on map rotation seems to provide a solid basis for further investigation. It remains to be seen to what extent people rotate interactive maps which are originally north-up into a state of being track-up, whether these rotations lead to more response accuracy and confidence, and which role time pressure plays in this context. This idea will be picked up again in Chapters 4 and 8, when experiments involving interactivity will be introduced.

2.5 User-related factors: Individual and group differences

When developing and evaluating geovisualization tools, the influence of user's individual and group differences is often not systematically assessed (Slocum et al., 2001). However, the discussion of previous work in the last sections has already demonstrated that many map-use studies have found significant interactions between map-related and user-related factors. In this section, I will emphasize the role user-related factors play for map-based decision making.

2.5.1 Terminology: Individual vs. group differences

As mentioned earlier in Chapter 1.2, one has to distinguish between individual differences and group differences when considering human influences on map-based decision making. The study of individual and group differences, also coined *differential psychology*, dates back to Charles Darwin in the 19th century. The underlying assumption of differential psychology is that each person is in certain respects either "like all other people" (i.e., the entire population), "like some other people" (i.e., member of a group of similar people), or "like no other person" (i.e., a unique individual) (Kluckhohn and Murray, 1953). Studies of individual differences (i.e., uniqueness) distinguish humans by parameters that can be measured on individuals, such as IQ scores, language abilities, or spatial abilities. Studies of group differences on the other hand, emphasize the aspect of difference across homogeneous groups of (similar) individuals (i.e., age groups, sex, or expertise).

2.5.2 Individual differences: Spatial abilities and their role in map-use studies

As Lloyd and Bunch express it, *"any study involving map use needs to take into consideration the differences among the individuals who use the map"* (Lloyd and Bunch 2010, p. 170).

The individual difference which has chiefly been investigated in connection with map-use studies is spatial ability. There are several ways and measures to assess humans' spatial abilities. Among these measures, the *Vandenberg Test Of Three-Dimensional Spatial Visualization* (Vandenberg and Kuse, 1978) can be regarded as the most frequently used one (Hegarty, 2010). In this paper-and-pencil-test, participants are asked to decide if block shapes match each other or not. It is widely accepted that most people solve the task by imagining rotating these shapes mentally. A sample task of this test is shown in Figure 9 in Chapter 4, when I will use this test within my empirical framework.

Hegarty (2010) provides a comprehensive overview of components and measurements of spatial intelligence and makes the link to virtual globes, mentioning the recent trend that *"children learn world geography by flying over the earth using Google Earth"* (p. 277), which has stimulated a new interest in spatial thinking and the acquisition of spatial intelligence. In this context, she also mentions the fact that not only instruction, but also playing video games can develop spatial intelligence to a certain degree. While males indeed seem to have significant advantages in mental rotation (Linn and Petersen, 1985; Voyer and Saunders, 2004), these gender differences can be reduced by playing video games for ten hours (Feng et al., 2007; Terlecki et al., 2008). In another study, Cohen and Hegarty (2007) have investigated correlations between spatial abilities and external visualization tasks. Discussing the use of 3D visualizations for learning anatomy, it has been shown that using 3D visualizations might be more problematic for low-spatial users (Hegarty et al., 2007).

Two competing hypotheses relating to the effect of spatial, or, more precisely, mental rotation abilities seem relevant for this study. Firstly, one might argue that because high-spatial participants are easily able to rotate objects such as maps in their heads, they would be less inclined to do it physically. This could also mean that high-spatial people would rotate a paper map less frequently during navigation in the real world, and thus might not need to use a map rotation tool in a digital and interactive map use setting either. In contrast, one could hypothesize that people with good internal visuo-spatial abilities are more likely to rotate complex external visualizations, especially those in 3D, as they are more likely to recognize the benefit of rotation for complex visuo-spatial tasks, compared to low-spatial participants. In fact, Cohen and Hegarty (2007) have shown that participants with good internal visualization abilities are more (and not less) likely to rotate 3D visualizations. This could mean that spatial abilities are indeed a necessary prerequisite for using an external visualization effectively and efficiently (see the work by Keehner et al. 2008). Results from the research by Lobben (2007) on paper maps mentioned above also favor the first hypothesis of a negative correlation between rotation abilities and actual map rotation.

2.5.3 Group differences in map-use studies: Sex, gender and experience

Group differences in cognitive abilities have been studied by psychologists in various fields, ranging from epidemiology to the assessment of linguistic skills. In studies about the efficiency and effectiveness of map-based decisions, group differences which have been investigated include sex (Weiss et al., 2003), gender (Lloyd and Bunch, 2008), age (Lloyd and Bunch, 2003), and previous training (Lloyd and Bunch, 2005).

For example, Weiss et al. (2003) have empirically confirmed the longstanding empirical finding that males have advantages in visuo-spatial abilities and map reading, whereas females have advantages in verbal tasks. Previous empirical studies do not show conclusive differences in performance between males and females in GIS- and map-related tasks (Albert and Golledge, 1999). This idea has been further investigated by Lloyd and Bunch (2005, 2008), who have studied individual differences in response time and accuracy with focuses on the variables gender and sex. They show that those variables have to be treated differently: Feminine gender had lowest response time and lowest mean accuracy, while "androgynous" persons performed fastest and most accurate. As for sex differences, females are reported to have advantages in memory of object locations, while males have significantly slower reaction times in map-reading tasks (Lloyd and Bunch, 2005). However, in a memory-location task, Lloyd and Bunch (2005) found female accuracy advantages and slower reaction times for males, and replicated results from previous studies regarding female advantages in acquiring spatial information from long-term memory (Birenbaum et al., 1994; Choi, 2003; Galea and Kimura, 1993).

Lloyd and Bunch have also taken into account other individual and group differences factors besides sex and spatial abilities, such as the working memory capacity, cognitive styles, the 2D/4D or digit ratio (Lloyd and Bunch, 2010), and right-handedness (Lloyd and Bunch, 2005), and shown that these factors might have similar or even more stronger effects on performance in certain map-related tasks than spatial abilities. Several authors have found that females take longer to complete wayfinding tasks (Delvin and Bernstein, 1997; Galea and Kimura, 1993), which might indicate that males might be the more efficient map-based decision-makers, that is, they perform relatively well under short time limits.

While males might be more confident and indeed perform better in certain map-related tasks, other research suggests a male overconfidence in map-based decisions. Various studies about the self-assessment of spatial intelligence (Furnham et al., 1999; Furnham, 2001) have demonstrated that males tend to overestimate their spatial abilities related to map-reading tasks, while females often underestimate them. This phenomenon has also been found for visual categorization with aerial photographs (Lloyd et al., 2002) wayfinding skills (Pedersen, 1999).

Several authors have shown that the ability to mentally rotate objects is one where sex differences are largest (Linn and Petersen, 1985; Voyer and Saunders, 2004). A variety of arguments exist for explaining this male advantage, such as environmental and socio-cultural differences, traditional gender roles (e.g., spatial tasks are perceived as being masculine in Western cultures), biological factors which emphasize differences in brain activity, but also spurious experiment-related variables, such as test time limits, task difficulty, or previous experience (see Parsons et al. (2004) for an extensive discussion). For instance, Prinzel and Freeman (1995) have uncovered that gender differences in spatial abilities increased with increasing task difficulty. Bunch and Lloyd (2006) argue that the male advantage in tasks involving the processing of visual images - like in the mental rotation test - might explain longer female response times in spatial tasks (Jordan et al., 2002; Stumpf, 1998).

Regarding map experience, Lloyd and Bunch (2005) found significantly higher accuracy and confidence values among geographers, but identical response times compared to non-geographers. These authors also showed that geographers were more successful than non-geographers in defining variations to categorize objects on aerial photographs (Lloyd et al., 2002). In a study on reading topographic maps, Chang et al. (1985) found that experienced readers performed better in interpreting contour lines. Albert and Golledge (1999) found no significant effects between GIS users and non-users in a GIS use study, and Soh and Smith-Jackson (2004) detected no experience differences in a wayfinding task. Overall, there is no clear tendency in favor of experienced GIS users in map-related tasks.

2.6 Summary

In the remainder of this chapter, I will identify the main issues for map-based decision making under time pressure emerging from the literature review and on which I will further focus in the experimental part of this thesis.

Firstly, research on decision making under time pressure has shown time pressure can be regarded as one of many factors that impair optimal human decision making. It is widely accepted that time pressure has a negative effect on the accuracy of and the confidence in human decisions for all sorts of tasks. This phenomenon is manifested in the concepts of the speed-accuracy and the speed-confidence trade-off. However, some researchers claim that the effect of time pressure resembles an inverted U-shaped curve, implying that time pressure could also have a positive effect on human decision making. An open question emerging for this study is whether map-based decisions follow a (perhaps linear) speed-accuracy or speed-confidence trade-off relation, or whether the patterns rather resemble an inverted U-shaped curve. At this point, it is also unclear how other factors, such as different levels of task complexity, different map types or users, might influence decision making processes under time pressure.

In map-based decision making studies, the element of time pressure has mainly been employed in navigation and route selection studies, domains where it is known for a long time that people have several “cognitive biases”, which prevent them from choosing the fastest route from A to B. However, time pressure has only rarely been used as an independent variable so far in this context, so that very little is known to what degree time pressure influences human road selection tasks or strategies.

Furthermore, it has been shown that different map types influence users’ preferences and the quality of their decisions. Common dichotomies regarding spatial displays in empirical map use research include the comparison of realistic vs. abstract maps or 2D vs. 3D. A frequent observation is that users generally prefer realistic or 3D maps to abstract or 2D maps, but actually do not perform better with these more realistic or higher-dimension displays. Indeed, there is a tendency that abstract representations (such as abstract 2D maps) seem to be more efficient for decision making than more realistic representations (such as 3D representations or satellite images), which suggests that abstract representations are more suitable for decision making under time pressure.

Issues of map design take into account the level of complexity or clutter on a map. Visual clutter seems to be a useful concept for this study on map-based decision making under time pressure, for it is directly measurable, and it can be assumed that visual clutter will have a negative effect on the efficiency of map-based decision making. An emerging hypothesis regarding map design for this thesis is therefore: The more clutter, the less efficient is a map for time-critical tasks.

In research about map interactivity, it has been demonstrated that users generally might overestimate the usefulness of interactivity, and the actual benefits of interactivity are still unclear. Therefore it seems interesting to investigate how human-map interaction influences the accuracy and confidence of user responses in specific tasks, especially as interactive, virtual maps have become ubiquitous as alternatives to static hardcopy paper maps.

As for user-related factors, several studies have found certain interactions between user characteristics (individual/group differences) and map-related factors (map design, spatial display types and interactivity). While some interaction tools (e.g., the map rotation tools by high-spatial participants) and spatial displays (e.g., realistic images by non-experienced map-users) may be more preferred by and more suitable for certain groups, other maps may be more efficient and effective for other groups (e.g., topographic maps for experienced map users). Sex and spatial abilities could be identified as the most relevant factors influencing the efficiency and effectiveness of map-based decisions. Nevertheless, still very little is known about how different individuals and groups cope with time pressure, and how map-based decisions under time pressure are actually influenced by the fact *who* is making the decision.

3. Methodological overview

In this chapter, I will present the empirical methods I chose to address the main research question (stated in section 1.3) of how time pressure might affect map-based decision making with different spatial display types and designs, interaction tools, by different people.

This chapter can be divided into two main parts: Firstly, I will describe the quantitative and qualitative empirical methods which I will employ for addressing this research question (sections 3.1 and 3.2). Secondly, I will present how I am going to measure the potential effect of time pressure and other variables on the efficiency and effectiveness of map-based decisions in four proposed experiments (sections 3.3 and 3.4).

3.1 Quantitative approach: Four controlled user experiments

The major part of my empirical approach consists of four controlled user experiments. All experiments share the recurrent theme of a typical emergency response situation, where people have to make a map-based decision under time pressure. Table 1 provides an overview of the experiments and their characteristics.

Table 1: Overview of the quantitative controlled user experiments.

	Task	Usage of time pressure	Independent variables	Map stimuli	Interactivity	Dependent variables/ measures
Exp I	Road selection in flat and hilly terrain	Binary variable (TP time pressure /NTP no time pressure)	Time pressure, map type, interactivity (within-subject), spatial abilities, sex (between-subject)	Six different map types	Four interaction tools	Suitability/preference ratings for maps and interaction tools
Exp II	Road selection in a flat urban environment	Three response time limits	Time pressure, display type (within-subject), task, sex (between-subject),	Two different display types (road maps and satellite images)	Not applicable	Accuracy and confidence
Exp III	Slope detection	Three response time limits	Time pressure, map type (within-subject), sex (between-subject)	Static 3D-looking maps	Not applicable	Accuracy and confidence
Exp IV	Several decision making tasks for the third dimension	Binary variable (TP/NTP)	Time pressure (within-subject), spatial abilities, video game experience, sex (between-subject)	Interactive virtual globe	Four interaction tools	Usage of interaction tools, accuracy and confidence

The aim of the first experiment (Experiment I) is to investigate map use (map type and interaction tool) preferences for decision making under time pressure in a road selection scenario. This is a typical map-based decision making task that can be performed both with and without time pressure. The results of this experiment can answer the question with which spatial display types and interaction tools users *believe* they would perform best with and without time pressure, and it can also answer the research question to what extent map use preferences depend on time pressure.

The display types which yield the highest suitability ratings in Experiment I require further investigation in follow-up experiments (Experiments II, III and IV). In these follow-up experiments, the display type and interaction preferences will be compared with actual map use performance (accuracy and confidence) for similar decision making tasks under time pressure.

Experiment II (Chapter 5) focuses on testing planimetric two-dimensional (2D) maps, which are suitable for road selection problems in which the third dimension is irrelevant. The road selection task in this experiment consists of selecting either the shortest route (in distance) or the fastest route (in driving time). As selecting the fastest route in driving time can be regarded as more complex than selecting the shortest route, one can also be assess whether the effect of time pressure might be more dependent on task complexity, and to what degree task complexity might influence the accuracy and confidence in spatio-temporal decisions. In this experiment, “task” emerges as an additional control variable. A between-subject design will be employed in order to control the variable task. I chose this design to eliminate the risk that participants confound the tasks, and also because the experiment time per participant is shorter compared to using a within-subject design (Martin, 2008).

In Experiment III (Chapter 7), a similar experimental setup will be applied to a scenario in which the third dimension is relevant for decision making. A special focus will be on the question of how display types that differ in their degrees of realism and clutter might influence map-based decision making under time pressure. The scenario of this experiment is informed by the results of prior expert interviews (see section 3.2).

The aim of the final experiment, Experiment IV, is to assess how interaction tool preferences measured in Experiment I are in accordance with how humans interact with a digital globe, specifically focusing on a 3D task (Chapter 8), and more precisely, if time pressure influences human-map interaction. Moreover, I want to investigate in this experiment whether time pressure influences the accuracy and confidence of map-based decisions with an interactive globe.

3.2 Qualitative approach: Expert interviews

To inform the design of the highly controlled experiments and to increase their validity (Barkley, 1991; Board, 1978), I carried out semi-structured guided interviews with several experts in the field of decision making under time pressure, such as helicopter pilots or ambulance drivers. The aim of these interviews is to shed light on the questions how (and which) maps are used by experts in the field and how typical time pressure conditions in map-based decision making actually look like, and to identify typical map use scenarios under time pressure which should be tested in future experimental setups. As these interviews were conducted after Experiment II and before designing and carrying out Experiment III, they are summarized in Chapter 6, in order be consistent with the chronological order of the empirical work.

3.3 Independent variables

In this section, I will present the four independent variables, whose effect on map-based decision making I am going to investigate.

3.3.1 Decision time limits

One main question of this thesis is how time pressure might influence the quality of map-based decisions. In order to measure the effect of time pressure, identical decision time limits for each participant will be used as independent variable in the experiments. These decision time limits are proxies for time pressure, as time pressure cannot directly be measured in a straightforward way within the scope of this thesis. Measuring individual stress levels or elementary information processes as indicators for time pressure is beyond the scope of this study.

The validity of the indicator of decision time limits is based on the assumption of a positive direct relationship between time limits and time pressure. In other words, I assume that *the shorter the time limits, the higher the time pressure.*

For two (Experiments I and IV) of the four experiments, time pressure will be used as a binary variable (*TP=time pressure* with a certain time limit, *NTP=no time pressure,* virtually unlimited decision time), while for the two other experiments (Experiments II and III) I will use three different decision time limits (and therefore three different time pressure levels). Pilot testing procedures will be conducted to validate the effects of the proposed time limits by ensuring that they lead to an increase in task difficulty and thus also to time pressure (see also section 2.2 for a discussion).

3.3.2 Spatial display types and map design

Another question is how the spatial display with which participants make decisions might influence the quality of their decisions. The spatial display types that will be used in Experiment I are orthorectified and oblique view satellite images, abstract road maps, topographic maps with contour lines, 2.5D and 3D representations (Wood et al., 2005), and shaded relief maps. The rationale for testing this multitude of spatial display types is that I want to investigate which kinds of maps users regard as suitable or unsuitable. The usage of other spatial display types for Experiment II will be dependent on the outcome of Experiment I and adapted to the road selection scenario, while for Experiment III I will also take the findings of the expert interviews into account. Regarding map design, I will assess to what extent visual clutter measures influence accuracy and confidence.

3.3.3 Interaction tools

Moreover, I want to explore how useful certain interaction tools and levels actually are for map-based decision making. The four current state-of-the art interaction tools which will be investigated are zooming, rotating, tilting and panning. Two of these tools relate to interactions with 2D maps (panning and zooming), while the other two tools (rotating and tilting) are common interactions with 3D maps.

3.3.4 User characteristics

Finally, one hypothesis is that different individuals will react differently to identical map stimuli and decision time limits. Among individual and group differences, the experiments will mainly investigate the individual difference of *spatial abilities* and the group difference of *sex*. Spatial abilities require special consideration when assessing the interaction with 3D globes and map rotation. While one can expect certain interaction effects between spatial abilities and sex (as seen in Chapter 2.5), the analysis of sex differences is also interesting for the discrepancy between confidence in and accuracy of map-based decisions, as it has been shown that males seem to be more confident in their decisions than females.

3.4 Dependent variables / Performance measures

In Experiment I, the preference or the suitability rating is the performance measure for a certain map type and interaction tools. There are no performance measures for the participants themselves. However, this changes for the three follow-up experiments II, III and

IV, where response accuracy and confidence are the two main performance measures and indicators for map use performance.

Accuracy is measured by the percentage of correct responses in Experiments II, III and one task in Experiment IV. In the latter, I will also measure accuracy by calculating the deviation from a certain correct value. *Confidence* involves the participants' retrospective self-evaluation of how strongly they believe that the decision they have made is correct. For the context of this thesis, I suggest that accuracy is the more important indicator, and that confidence is only of secondary importance. The underlying assumption for this is that confidence cannot be a valid performance measure when the decision was not accurate. However, if participants have given equally accurate answers with two different map types, the map which also leads to a high confidence in the decision can be regarded as the more successful one for decision making tasks. As participants always have to make a decision before proceeding to the next task in the experiments, it is important to measure whether they felt confident about their choice or just chose a route randomly because they were asked to do so. High confidence ratings indicate that participant responses are not necessarily the result of guessing, but that the maps have actively supported participants in their decision-making processes.

These performance measures of accuracy and confidence, which are originally assigned to one map stimulus, can then be aggregated to create overall average values, for instance, per participant, per group (e.g., male/female participant), per time limit, or per map type.

As mentioned in section 1.2, *effectiveness measures* assess to what extent people give right answers with a map. Response accuracy and confidence are primarily indicators for effectiveness. In other words, a map-based decision is *effective* if it is accurate and if the decision-makers are also confident in their accuracy. In contrast, a map-based decision is more *efficient* if effectiveness measures are also high under time pressure, or short response time limits.

In Experiment IV, I will also measure human-map interactions, more precisely, the usage of certain map interaction tools. While interaction tool usage is not really a "performance measure" as such, the aim of measuring interactivity is to analyze whether people who interact more with a map are also more accurate solving complex tasks and confident in their decisions. Finally, it should be noted that response time will be used as an independent control variable throughout the study. Unlike in previous studies, response time is thus not a performance measure within the scope of this study.

4. Experiment I: Map use preferences with and without time pressure

The literature review in Chapter 2 has shown that road selection is a very common task for map-based decision making, which can be performed under both time pressure and no-time-pressure conditions. Under time pressure (TP), a classic scenario might be to reach a certain place as quickly as possible (emergency situation). In contrast, when planning a holiday trip well in advance, route selections will not be driven by severe time constraints (no time pressure = NTP). For this purpose, a map-based route selection task was chosen for Experiment I, which is described in this chapter. Participants were exposed to an emergency response (time pressure, TP) scenario and an excursion planning (no time pressure, NTP) scenario.

Moreover, previous work has suggested that map use preferences are often not in accordance with map use performance, which is frequently measured in accuracy and confidence. In particular, the benefits of realistic representations and interactivity are often overestimated. Therefore, one goal of this experiment was to explore how time pressure influences map users' preferences for display types and interaction tools in a road selection task, dependent on whether they had to select routes under time pressure (emergency situation) or not (trip planning), before measuring actual map use performance (accuracy and confidence) in Experiments II – IV. In this experiment, map use preferences were operationalized by two indicators: spatial display preferences and interaction tool preferences. Regarding interactivity, one focus was on the question whether zooming and panning are also considered the most important tools under time pressure, as Harrower and Sheesley (2005) suggested.

As mentioned in section 1.3, another research question for this work is to what extent map display and interactivity preferences depend on individual and group differences. As several researchers (Linn and Petersen, 1985; Voyer and Saunders, 2004; Weiss et al., 2003) have found a relationship between gender and mental rotation abilities, or, more precisely, that male participants seem to have advantages at visuo-spatial tasks such as mental rotation, one might also expect potential interactions of these two factors when assessing the effectiveness and efficiency of map-based decision making. A hypothesis in this context is that male map use preferences are more similar to the preferences and performance of so-called "high-spatial" people (i.e., those participants who would typically score high on a Mental Rotation Test), and female preferences are similar to the preferences of "low-spatial" people.

As discussed in section 2.4.2, the influence of spatial abilities on the preference for the map rotation tool is of particular interest in this context. The experiment reported in this chapter can answer the debated question whether high-spatial participants have a higher or rather a lower preference for using the map rotation tool.

4.1 Participants

One hundred fifty-five participants (104 male, 51 female) took part in this experiment. Seventy of those 155 participants were recruited from the local university, mainly from geography and cartography courses. The remaining 85 participants were recruited via an international cartography mailing list. Most of the participants were experts in GIS or cartography, as 79% of these participants reported cartography, and 89% GIS experience of more than one year. Ninety-seven percent reported they used maps at least occasionally in their leisure times.

The intended population was the person whose everyday job is to make map-based decisions, but not necessarily always under time pressure. It can be assumed that this person will be familiar with maps and their use. Thereby, a population sample with varying backgrounds was chosen, of which a great share had considerable map use experience.

4.2 Materials

Six map types with different designs were chosen as stimuli for this experiment (shown in Figure 7), depicting different mountainous areas from all over the world. The test maps all had identical sizes (145 x 189 pixels). As mentioned in the previous chapter, the distinction between spatial display types is not based on varying display sizes (e.g., PDA vs. large screen) within this thesis, but is based on different ways of representing spatial phenomena on a desktop computer screen.

The following map types were chosen because they represent typical and commonly encountered map types for road selection tasks:

- Map A: Terrain map with shaded relief
- Map B: Topographic map including contour lines
- Map C: Road map including shaded relief and spot heights
- Map D: Satellite image with oblique 3D perspective and elevation exaggeration
- Map E: Road map including labels with distance information
- Map F: Satellite image with orthogonal perspective including roads and map labels

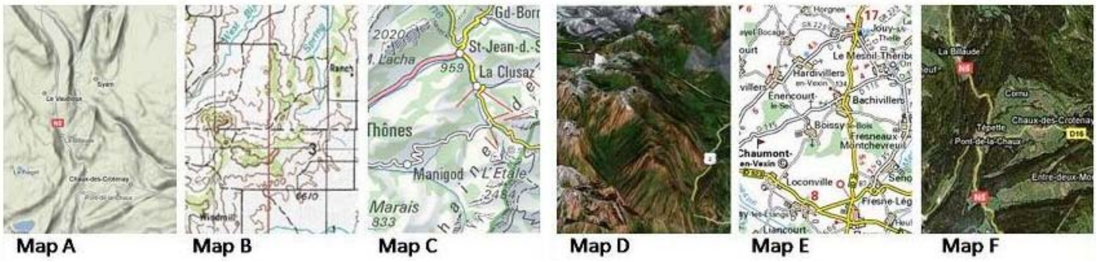


Figure 7: Six spatial display types participants were asked to rate according to task suitability.

As mentioned before, a variety of different map types was tested to get first insights into spatial display preferences and especially in order to identify the most suitable display types.

For testing interaction tool preferences, I created screenshots of the four most common interaction tools (see section 2.4.1) and added static 2D images, which show the map view before and after using the respective interaction tool. Additionally, I included verbal descriptions of the tools and their function below the images, as shown in Figure 8.

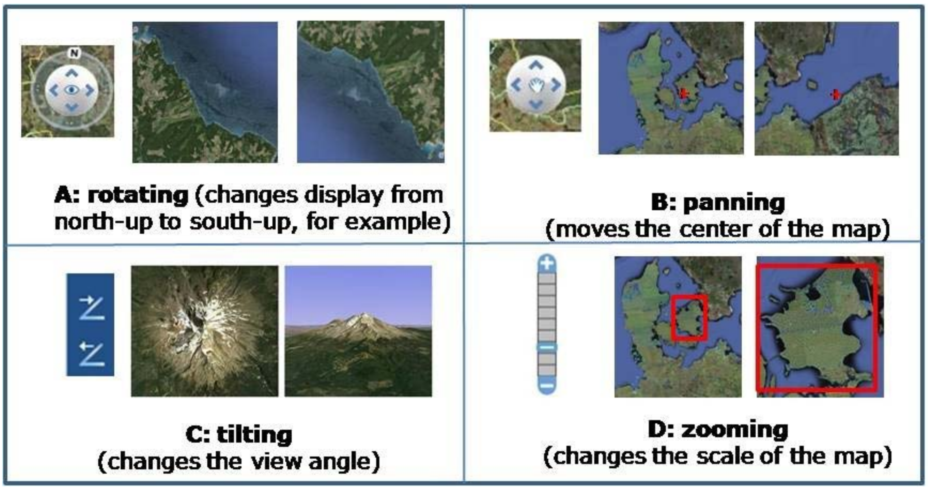


Figure 8: Stimuli describing the four interaction tools, which participants were asked to rate according to task suitability.

4.3 Procedure

For the seventy participants who were recruited from the local university, the experiments took place in a controlled environment, that is, in a lab equipped with standard personal computers connected to the Internet. These participants first performed a paper-and-pencil version of the Vandenberg's Mental Rotation Test (Vandenberg and Kuse, 1978), because, as mentioned earlier, I was also interested in identifying if and how spatial abilities might have an influence on user preferences. A sample task of the Mental Rotation Test (MRT) is shown in Figure 9. The mental rotation test consisted of 20 similar mental rotation tasks. After the instructions, participants were given six minutes to complete the entire test.

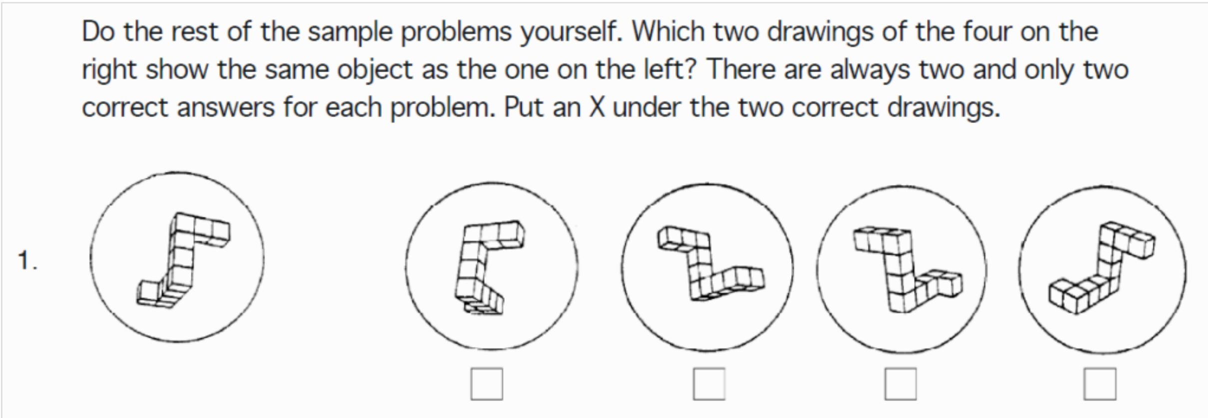


Figure 9: Excerpt from the Vandenberg Test of 3D Spatial Visualization (Vandenberg and Kuse, 1978).

The main experiment was conducted on a standard web browser. Participants were first asked to fill out a background questionnaire, and then they were introduced to the two map use scenarios. The time-based decision making scenario (TP) included saving a friend as quickly as possible, while the scenario without time-pressure (NTP) involved excursion planning under no time constraints. The sequence of the two scenarios was systematically alternated, so that roughly half of the participants started with the emergency response (TP) scenario including an animation counting down the time, and the other half started with the excursion planning (NTP) scenario.

Time pressure was simulated with a count-down animation starting at 60 seconds and counting backwards until a red colored text box appeared when time was up, blinking at high frequency and including the phrase: "Please answer now". This 60 second limit was chosen after pilot testing. As some of the pilot test users were not able to answer the question within a shorter time limit, I selected the 60 seconds threshold for this study.

For both scenarios, participants were asked to indicate their preferences for the series of display types (Figure 7) and map interaction tools (Figure 8). The order of the display types and interaction tools on the page within a scenario was also randomized to prevent any potential ordering biases. For each map and task, participants were first asked to indicate their spatial display preferences on a rating scale ranging from "1 – the map is not suitable" to "5 – the map is very suitable" for both the TP and NTP scenario. In the second portion of both the TP and NTP scenario, participants had to indicate their interaction tool preferences for the task on a rating scale ranging from "1 – I would definitely not use it" to "5 – I would definitely use it". After completing each rating task, users were asked to click a "Next Page" button to proceed to the next task.

The last portion of the test included open-ended questions. Participants were asked to provide explanations for their ratings, and whether they felt that they had enough time for solving the task. The duration of the web experiment was approximately 15 minutes.

The mailing list participants did not take the MRT before and could start the test any time after receiving an e-mail with the respective URL of the experiment web page. They were instructed to solve the task on a computer screen with a resolution of at least 1280 x 768 pixels.

4.4 Results

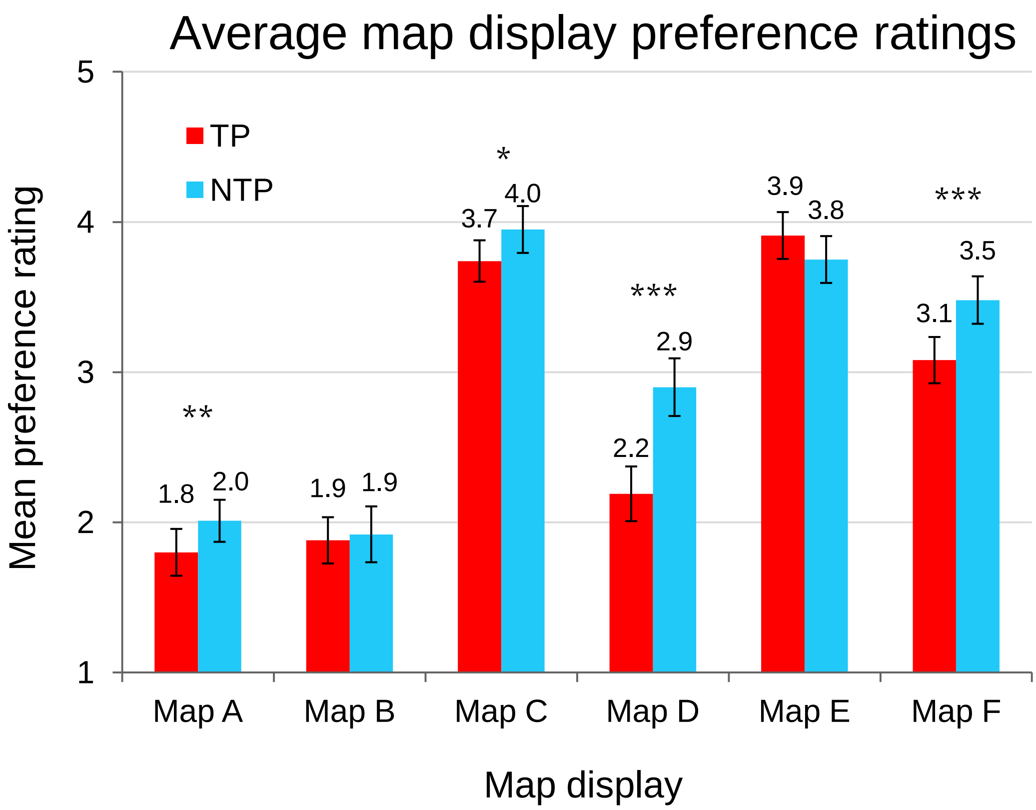
First, I will discuss participants' quantitative preference ratings on map types. This is followed by the results on the interaction tools. Finally, I will summarize how spatial abilities and sex related to the preferences for map displays and interaction tools.

4.4.1 Map type preferences

The summary of the map display ratings is shown in Figure 10. The road map with shaded relief and explicit elevation information (Map C) obtained the highest ratings for the excursion planning (NTP) task ($M=4.0$, $SD=0.9$), while the road map without shaded relief, but including explicit distance information (Map E), was most preferred for the emergency situation (TP) task ($M=3.9$, $SD=1.0$). As for the worst ratings, the topographic map (Map B) received the lowest rating ($M=1.9$, $SD=1.0$) in the NTP condition, while the terrain map (Map A) was least preferred ($M=1.8$, $SD=0.9$) for the TP task. Both satellite images (Maps D and F) yielded a rating close to 3 (basically the midpoint of the rating scale with no clear preference).

The satellite image with oblique 3D perspective (Map D) has the highest variance in the preference ratings for both tasks (SD TP=1.1, SD NTP=1.2), which indicates that there is least agreement on its task suitability under both temporal conditions.

The most realistic looking satellite maps (D and F) also obtained considerably lower suitability ratings under time pressure compared to all other maps. In an open answer at the end of the experiment, more than two thirds of the users (67.1%) specified they had rated the maps differently under the two conditions (see also section 4.6). As the rating results were not normally distributed, a Wilcoxon Signed Ranks Test was conducted to test whether the mean ratings for display types differ under the two scenarios. This analysis suggests that all maps except Map B (the topographic map with contour lines) and Map E obtained significantly different ratings under the two tested scenarios ($p < .05$). In other words, participants preferred different maps depending on time pressure.



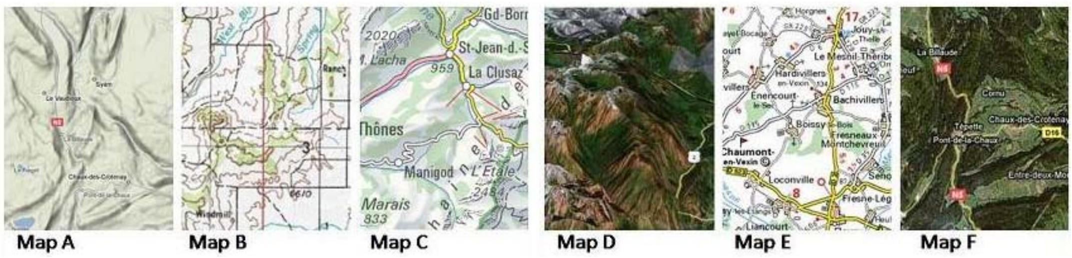


**Figure 10: Average preference ratings for different map display types. A rating of 5 means "The map is very suitable", while a rating of 1 means "The map is not suitable".
Error Bars: ± 2 SE (Standard Error). *** p< .001, ** p< .01, * p< .05.**

Another way to look at this is to evaluate whether the maps were rated significantly higher or lower in suitability than the midpoint of the rating scale (3), indicating no clear preference. A one sample t-test suggests that rating results differ significantly from the mid-point of the rating scale, with the exception of the 3D satellite map (Map D) for the NTP condition, and the 2D satellite map (Map F) for the TP condition (p< .001). Hence, aside from these two mentioned maps, all others were either clearly rated as suitable or unsuitable for the tasks at hand.

4.4.2 Interaction tools

The ratings of the interaction tools are summarized in Figure 11. The zooming tool was preferred most for both the NTP and TP scenarios. Under both conditions, the vast majority of the participants (135 out of 155; 87.1%) indicated that they would definitely use a zoom tool under time pressure and under the NTP scenario (M=4.8, SD=0.5 for TP, M=4.9, SD=0.9 for NTP). Panning obtained the second-highest rating (M=4.5, SD=0.9 for TP, M=4.6, SD=0.8 for NTP), followed by tilting under both conditions. Rotating was regarded as least important for both the NTP (M=2.6, SD=1.3) and TP (M=2.6, SD=1.4) tasks.

There was least agreement on the usefulness of the rotation tool under both conditions (SD TP= 1.3 and SD NTP=1.4), irrespective of time pressure. As Figure 11 shows, the tilting tool was assigned significantly lower ratings in the TP task (M=2.7) compared to the NTP task (M=3.4). The ratings for tilting ($p < .001$) and panning ($p < .05$) differ significantly between the two scenarios, while the ratings for zooming and rotating did not differ at all ($p > .05$).

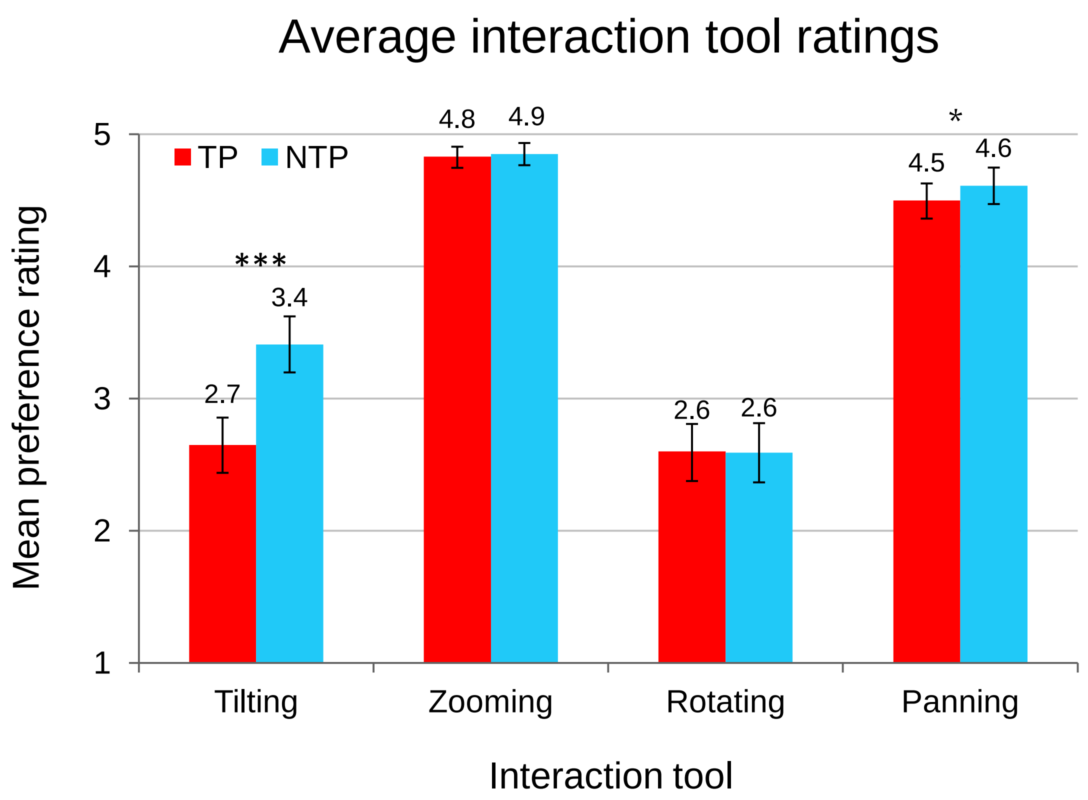


Figure 11: Average preference ratings for different interaction tools. The scale ranges from 1 ("I would definitely not use it") to 5 ("I would definitely use it"). Error Bars: ± 2 SE. *** p< .001, * p< .05.

Furthermore, a one sample t-test suggests that all ratings for the interaction tools differ significantly from the mid-point of the rating scale (3), suggesting that participants are not indifferent about tool suitability. More than sixty percent of the users (62.6%) stated that they had rated the interaction tools differently for both tasks.

4.4.3 Spatial abilities and their influence on map use preferences

On average, the 70 participants who participated in the MRT scored 20.8 points (SD=8.5) out of 40 possible points on the Mental Rotation Test (MRT). The median score was 20.0 points. The male (N=40) average score was 23.4 (SD=8.2), while females (N=30) scored 17.4 (SD=7.8) points on average. The boxplot in Figure 12 shows the response distribution by sex in graphic form. An independent samples t-test confirms that MRT scores grouped by sex are indeed significantly different ($p < .01$). In other words, the participants represent the expected sex differences in mental rotation abilities mentioned in section 2.5.2 (Linn and Petersen, 1985; Voyer and Saunders, 2004).

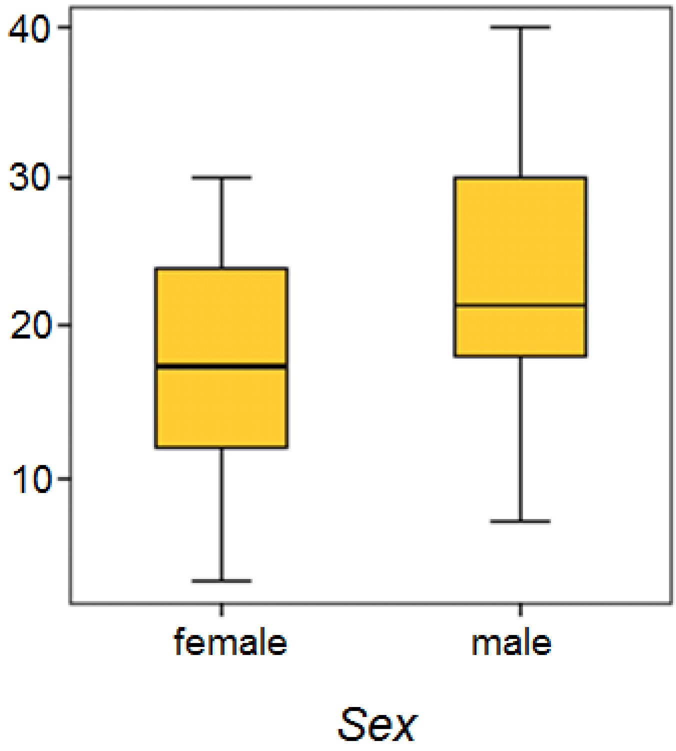


Figure 12: Boxplot of the MRT scores by sex.

For analyzing how mental rotation scores influence participants' preferences for map display types and map interaction tools, I assigned the 70 test participants who had participated in the MRT to two groups by a median split on the MRT score, as has been done in similar studies that did not include map stimuli (Downing et al., 2005). Thirty-three participants thus were assigned to the "low-spatial" group and thirty-seven to the "high-spatial" group, respectively.

Mental rotation abilities did not significantly influence participants' map type preferences when using the median split. However, some effects of mental rotation abilities on interaction tool preferences are evident: On average, except for the tilting tool, the high-spatial group showed a higher preference for three out of four interaction tools (i.e., zooming, rotating and panning) compared to the low-spatial group, regardless of the time pressure context (see Figure 13).

The map rotation tool was preferred more by the high spatial group than the low-spatial group for both TP and NTP scenarios. The mean preference ratings under time pressure are 2.8 (SD=1.6) for high-spatial participants and 2.2 (SD=1.0) for low-spatial participants. Without time pressure, the average preference ratings are 2.8 for the high spatial (SD=1.4) group and 2.3 (SD=1.2) for the low-spatial group. While the differences show the expected tendencies, they are, however, not statistically significant.

Mental rotation abilities seem indeed to significantly affect the preferences for the 2D interaction tools: Participants with high spatial abilities (N=37) have a significantly higher preference for panning (M=4.7, SD=0.5) under time pressure, compared to the low-spatial (N=33) group (M=4.2, SD=1.1). Similarly, the high-spatial group has a significantly higher preference for zooming (M=5.0, SD=0.2) under time pressure than the low-spatial participants (M=4.8, SD=0.6).

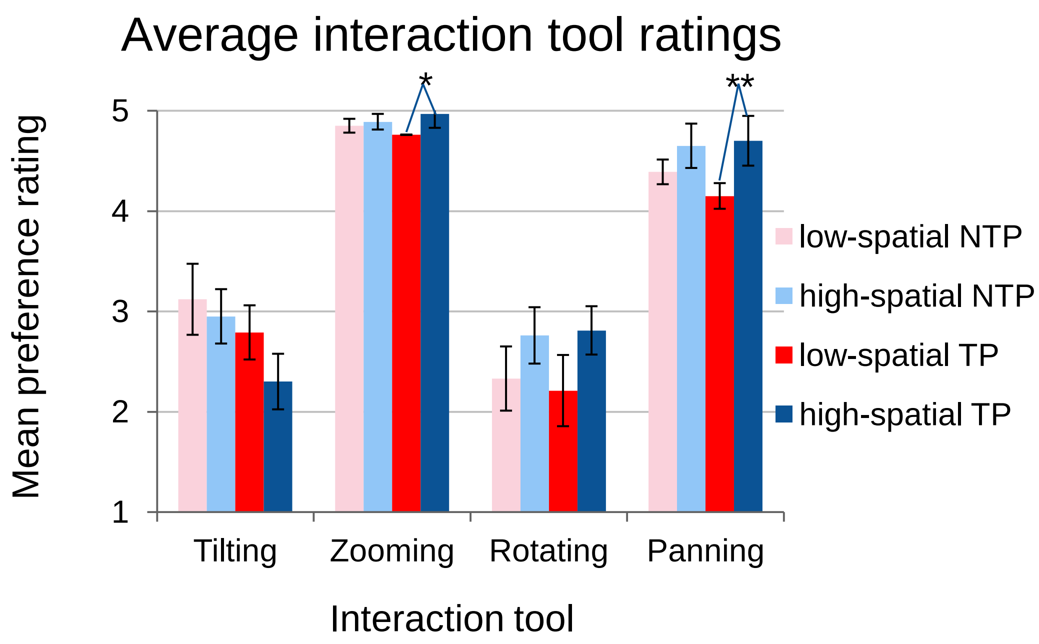


Figure 13: Interaction tool preferences dependent on time pressure and spatial abilities, using two subgroups. Error Bars: ± 2 SE, ** p< .01, * p< .05.

Similarly to what Cohen and Hegarty (2007) have found in a 3D rotation and perspective taking task, the high-spatial participants not only seem to have an overall higher preference for using interaction tools, but also specifically prefer being able to rotate a map for solving a road selection task, irrespective of the time pressure context.

4.4.4 Spatial abilities analysis with three subgroups

Next, I wanted to explore the robustness of these findings considering the median-split aggregation procedure of individual MRT scores. Following Sholl and Liben (1995)'s approach, I divided participants into three spatial ability groups using terciles of their MRT scores and assigned them to a low-spatial (N=24), medium-spatial (N=26), and high-spatial group (N=20). In the following, I present and discuss the results only for the newly created high-spatial and low-spatial groups, and do not further consider the medium-spatial group.

Comparing Figures 13 and 14, one can see that the overall response pattern is almost identical. In other words, regardless of the high-low spatial grouping procedure, one can find the same preference patterns for the map interaction tools. However, while the lower preference for the zooming tool in low spatial participants was significantly different from the high-spatial group when using a median-split grouping, this significant difference disappears using the tercile approach. In contrast, the significantly lower preference for the panning tool in low spatial participants remains significant ($p < .01$), regardless of the classification scheme.

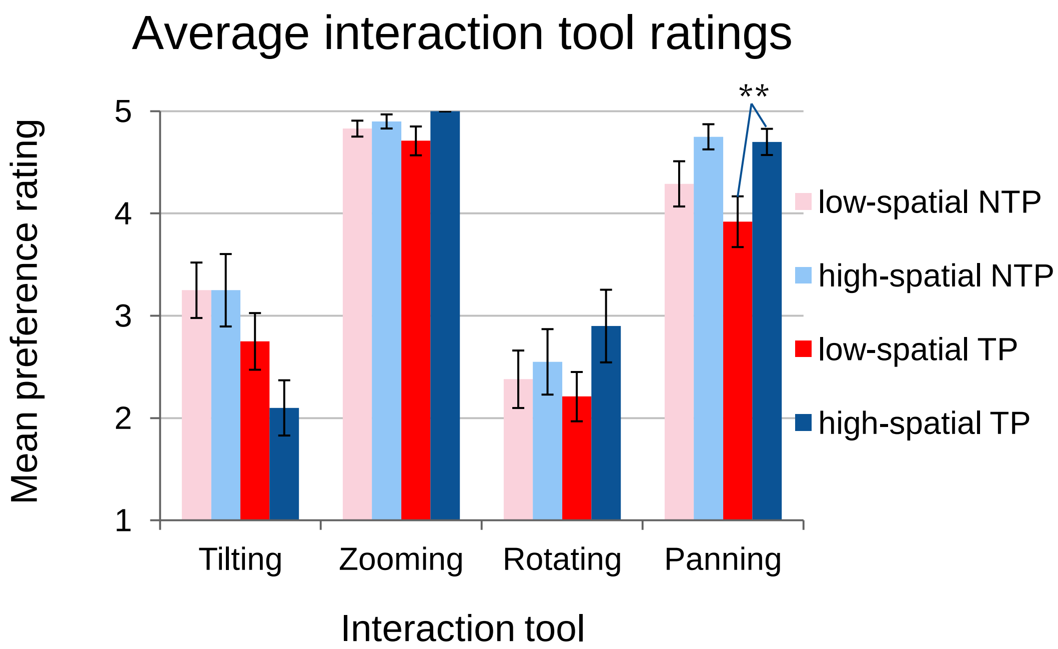


Figure 14: Interaction tool preferences when using three classes of spatial abilities. Preferences for the medium-spatial group are not displayed in this figure. Error Bars: ± 2 SE, ** p< .01.

When analyzing the effect of these newly built groups on map display preferences, new significant effects emerge: For the excursion planning scenario (NTP), the low-spatial group has a significantly higher preference for Map C (road map with shaded relief). As for the emergency response (TP) scenario, the preference for Map E (the road map with distances) is significantly higher for the high-spatial group (see Figure 15). This is in contrast to the median split classification with two groups, where there were no significant differences in map display preferences.

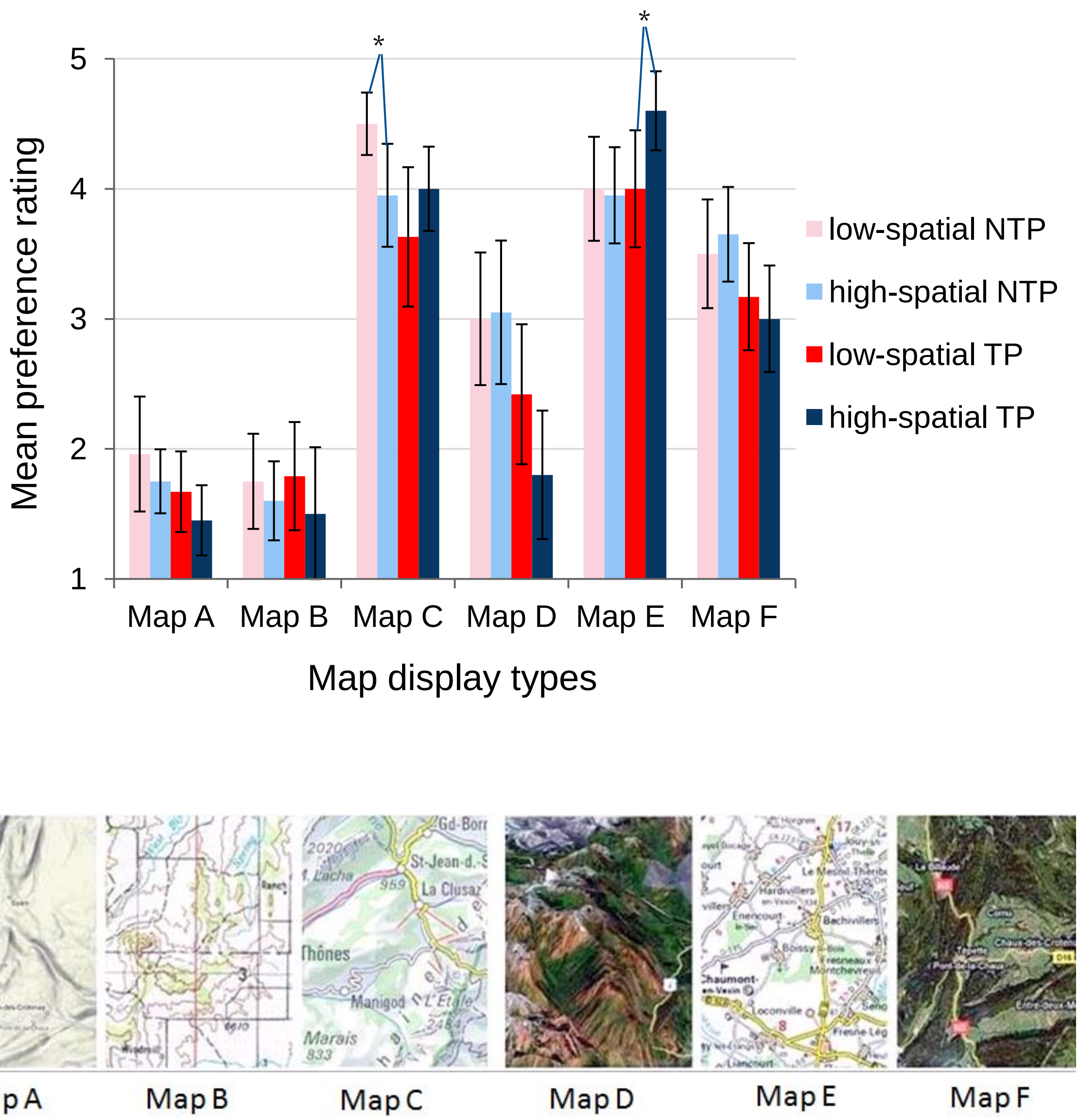


Figure 15: Map preferences according to classification into three classes of spatial abilities. Error Bars: ± 2 SE, * p< .05.

These results show that possible interpretations about the effect of spatial abilities on map display and interactivity preferences might significantly depend on the classification methods used to study these individual differences. This further means that researchers have to take great care in validating found effects by cross-checking results with various established methods.

4.4.5 Effect of sex on spatial abilities

As reviewed in the related work section earlier, sex and mental rotation scores are interdependent. One might therefore expect that on average more males will be found in the high spatial group, and thus have similar visuo-spatial preferences, while female preferences will be more like those of the low-spatial group.

A Chi-Square test of independence assessing the relationship between MRT response patterns (grouped by a median-split) and sex suggests a significant relationship between these two factors (χ^2=5.85, $p< .01$). This relationship remains significant when using the high-low spatial groups from the tercile classification (excluding the medium-spatial group). However, with the tercile classification, the association is considerably weaker (χ^2=3.88, $p< .05$).

Even though there seems to be a significant dependence of sex and spatial abilities, the male participants do not have the same map use preferences as high-spatial participants. Similarly, the female participants do not share the same map use preferences of low-spatial participants, as will be seen in the next section.

4.4.6 Effect of sex on map use preferences

First and most importantly, there were no significant sex differences in map interaction tool preferences. While the overall map type preference patterns are strikingly similar in Figures 15 and 16, one can see that male/high-spatial and female/low-spatial patterns are not congruent. I found highly significant sex differences in map type preferences: Map E (the road map including road distances) was significantly more strongly preferred by females (NTP: M=4.1, SD=0.9, TP: M=4.4, SD=0.7) than by males (NTP: M=3.6, SD=1.0, TP: M=3.7, SD=1.0) in both conditions. This is particularly interesting as high- (and not low-) spatial people preferred Map E significantly more than the other maps in the same condition (see Figure 15). Female preferences are also higher for the other road map, C, under both conditions, and this difference is significant for the NTP condition (see Figure 16).

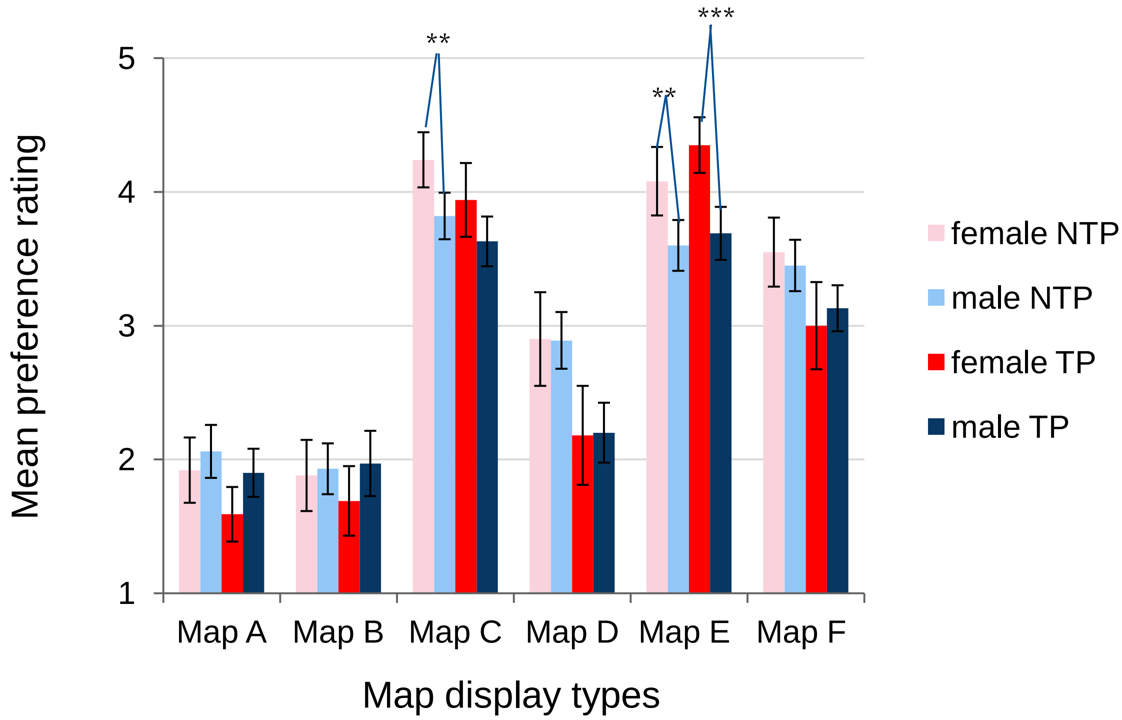


Figure 16: Map display preferences dependent on time pressure and sex.
Error Bars: ± 2 SE, *** p< .001, ** p< .01.

4.5 Analysis of open-ended answers

I also collected users' self-reports rating explanations collected in open-ended questions. The aim of asking these open-ended questions was getting further insights about possible reasons that could explain users' preference ratings. I will summarize users' self-reports in the next paragraphs.

Under time pressure, participants need to specifically find the relevant information about how to get from point A to B. Aesthetic details and information irrelevant to the task at hand are seen as less important *("I don't need super fancy view when I'm in rush. I rather need route classes to find my way faster", "The maps which are optically nice do not contain more information regarding accessibility of a place", "If I can take my time I would bear in mind other* *things like panorama")* under this condition. Some participants assigned their highest rating to the road map (Map E) because it was the "least fancy one", while others stated that the top-down 2D perspective was good for estimating distances. A number of participants mentioned they chose road maps over 3D displays for the TP task because they were used to them. Overall, under time pressure, users seem to prefer maps with which they are more familiar compared to the no time pressure situations.

Moreover, people believe they need more information about the relief in the excursion planning (NTP) condition, but more information about the road network in the TP condition. General accessibility of locations, particularly the road network, and precise locational information to identify a place are regarded as vital under time pressure. While the saliency of roads and a road network classification seem to be the most important factors for the TP task *("without roads or tracks it's unusable")*, participants prefer a good overview, shaded relief and other aesthetic details of the natural landscape for the excursion planning (NTP) scenario.

Summarizing the findings for the map types, one can conclude that for the NTP task, natural landscape aspects and overall impressions of what to see and to expect appear to be important. On the other hand, precise locational information for more accurately determining the shortest/fastest path along a road network including quick and easy access to task relevant information are essential under a time pressure scenario.

The main findings about the suitability of the interaction tools can be summarized by the following participant statement*: "Zooming and panning are important. The other two are just for decorative purposes."* That is, under time pressure, users seem to require zooming and panning as essential tools, while the 3D interaction tools, tilting and rotating, are regarded as superfluous. Tilting was often regarded as "fun" and unnecessary under time pressure *("when no time, tilting should not be used"* or *"tilting and rotating are rather toys for me")*, which explains the relatively large difference in ratings between the two conditions for this tool.

In contrast, some users even assigned higher ratings to all the interaction tools under time pressure compared to the NTP scenario. Some participants stated that they would *"probably maximize all the tools available in an emergency",* while others argued that they needed to exactly locate a person, so that they would need more information which they would only get from using interaction tools. Finally, most participants stated that their rating of maps and interaction tools would depend on additional, more detailed information about the excursion or the emergency situation.

After analyzing user preference ratings and their potential explanations, different preferences for display types and interaction tools under time pressure (TP) and no-time pressure (NTP) conditions can be summarized in Table 2 as follows.

Table 2: Requirements for maps under different temporal conditions.

Excursion planning / no time pressure (NTP)	**Emergency rescue / time pressure (TP)**
• information about the natural landscape, terrain • hill shading, relief • good overview • aesthetic details • interaction tools to explore the area with various views • tools that are fun to use	• quick access to information to precisely locate a person • details about accessibility / road network • good saliency of roads • classification of roads • planimetric view for cartometric tasks • clear and unambiguous display • familiar map styles • efficient tools that quickly render an exact "best view"

4.6 Experiment I – Discussion

The goal of this first experiment was to answer the research questions (as stated in section 1.3) whether people's preferences for map display types and map interaction tools depend on time pressure. The results suggest that these questions can be answered in the positive, as participants indeed have different map display and interaction tool preferences when using maps for decision making, depending on whether they are under time pressure or not.

Time pressure particularly affects the preferences for 3D maps and the respective interaction tools: Both realistic satellite display types and the tilting tool are significantly less preferred under time pressure. Possible explanations for these differences are that the satellite maps contain on the one hand more aesthetic details, which are not important under time pressure, but on the other hand less information about road networks, which are essential under time pressure. As for tilting, users consider this as an exploratory tool which is not useful under time pressure.

Users were also asked whether time pressure influenced their map and interactivity ratings. More than two thirds of the users stated they had rated the maps differently, while more than 60% claimed they had rated the interaction tools differently under both conditions. Statistical analysis supports users' beliefs that preference ratings for the map displays and (to a lesser extent) for the interaction tools do indeed differ significantly across conditions. When under time pressure, users seem to prefer quick access to precise, task relevant information, including familiar map displays showing mainly a well identifiable road network. Without time pressure, participants preferred 3D, landscape details and aesthetic aspects, including a good overview, and interaction tools to playfully explore the map. However, the overall strong preferences for using the familiar road maps and the 2D zooming and panning interaction tools seem to be robust and not affected by the time pressure condition. The high suitability rankings for panning support the view of Harrower and Sheesley (2005) that panning and zooming are the two most important interaction tools.

In general, the preferences for map display types seem to be slightly more affected by time pressure than the interaction tools, and static 2D representations seem to be most adequate for time pressure conditions. This finding is in accordance with several previous studies (Coors et al., 2005; Dillemuth, 2005; Smallman et al., 2001) discussed in section 2.3.1, which have found that static abstract 2D representations seem to be more suitable for time-critical map-based decision making than interactive realistic 3D representations.

User-related factors

While I could replicate an interaction effect of sex on spatial abilities, confirming results from previous research (Linn and Petersen, 1985; Voyer and Saunders, 2004), in this experiment sex does not influence map type and interaction tool preferences in the same way that spatial abilities do. As different classification methods lead to different significant preferences, these results also suggest that the methods employed to measure and group participants based on spatial abilities can influence the robustness of empirical research results.

In this context, one also has to bear in mind two issues regarding sample sizes: Firstly, the participant number in this experiment is relatively high (N=155). Therefore, the chances of a "false positive" or Type I errors increase. In other words, the large sample size might explain that the effect is statistically significant, while in reality there is no sex difference in map preferences. Secondly, the sample sizes of male and female participants are unequal. This would be a major problem if the smaller of the two sample sizes was particularly small, or if the variances were inhomogeneous. However, as discussed above, none of this is the case in our experiment. In contrast, the large sample sizes result in very high values of statistical power for the significance tests, that is, the chances of a "false negative" or Type II error are very low (Fields, 2009).

In general, map type preferences (such as higher preferences for road maps among female participant) seem to be stronger affected by sex, while interaction tool preferences are more influenced by spatial abilities. Perhaps somewhat counter-intuitively, my results show that high-spatial participants overall have higher preferences for map interaction tools than low-spatial participants. Good mental rotators also prefer to use the map rotation tool more than low-spatial participants, but this difference is not significant. This is in accordance with findings by Cohen and Hegarty (2007), who have found that participants with good internal visual abilities are more (and not less) likely to rotate external 3D visualizations. Altogether, my results indicate that user characteristics have indeed a strong effect on map use preferences. The fact whether participants conducted the experiment in a controlled condition (N=70) or via the internet (N=85) had no significant effect on preference ratings.

Limitations of the map use preference experiment

While this first experiment has generated initial useful insights on people's map use preferences under time pressure, one has to bear in mind certain caveats when generalizing the results. Firstly, one single time limit for all users was employed. However, there are a many more factors to consider than simply limiting task completion time, as mentioned in section 2.1.2. The fact that an identical response time limit already leads to different perceptions of time pressure is also reflected in the fact that 78.7% of the participants felt that they had enough time to rate all the maps, while 21.3% wished they had been given more time than 60 seconds. For these 78.7% who claimed to have had "enough time", it does not necessarily mean that they were not under time pressure. This just indicates that they had enough time to complete the task. Although setting a shorter time limit would have arguably increased the time pressure, a resulting shortcoming of using a shorter time limit would be that a number of participants would perhaps not have been able to complete the rating task, which I tried to avoid for this study.

Moreover, the ratings for the map display types will always be influenced by the specific map. In this experiment, the poor ratings for the topographical map could also be a result of poor design, where another, more aesthetically pleasing topographical map could have yielded higher ratings. This is especially relevant in the excursion planning scenario, where participants indicated that aesthetic details did matter for their rating decisions.

Finally, the task in the excursion planning scenario was rather vague. It is not clear whether participant ratings would have been different with more concrete tasks in the non-time-pressure condition, such as the means of transport or the type of excursion. Moreover, differences in preference ratings might not only be due to time pressure, but also due to the differences in the scenarios (emergency response vs. excursion planning). While one can assume that these two tasks represented typical time pressure and no-time pressure scenarios respectively, users might also have different map and tool preferences which are highly task-specific, and not time-pressure-specific.

Implications for follow-up experiments

Experiment I has chiefly shown that ratings for road maps and satellite images were generally higher than for topographic maps and shaded relief maps. Therefore, road maps and satellite images should be further investigated in a follow-up road selection experiment, where people's actual task performance with these display types can be measured and compared with the preference ratings from Experiment I.

As participants preferred the more 2D abstract road maps over the more realistic 3D looking satellite maps for both TP (time pressure) and NTP (no-time pressure) conditions, an open question is whether people also give more accurate answers with road maps than with satellite images.

Another research question arising from Experiment I is whether map users actually perform better with satellite images when there is less time pressure, as the preference ratings from Experiment I would suggest.

5. EXPERIMENT II: ROAD SELECTION WITH STATIC 2D MAPS

One key finding of Experiment I was that time pressure and map types influence human preference judgments. The main goal of Experiment II was to investigate whether time pressure and display types also affect user performance, that is, accuracy and confidence, for a road selection task (with and without time pressure). In this experiment, time pressure was manipulated by including three different response time limits that had been identified in pre-testing: a severe time limit (10 seconds), a moderate time limit (20 seconds) and a generous time limit (30 seconds).

I chose two spatial display types that had obtained the highest suitability ratings in Experiment I: 2D orthographic satellite images and classic 2D road maps. Using these two display types, I wanted to tackle the research question (see section 1.3) of how spatial inference- and decision making under time pressure might depend on the spatial display type and design. Furthermore, this experiment can shed light on a question regarding road maps and satellite images that emerged from Experiment I: Firstly, do people generally make better map-based decisions in road selection tasks with the more preferred road maps than with satellite images? Secondly, how does time pressure influence users' decision making with satellite images?

As discussed in Chapter 3, response accuracy and confidence are the dependent variables in my controlled experiments. Measuring these variables, I want to examine to what extent road selection choices under time pressure resemble the patterns of a speed-accuracy trade-off (see chapter 2.1.3), a speed-confidence trade-off (see chapter 2.1.5), or rather an inverted U-shaped curve (see chapter 2.1.4). Previous work has shown that the speed-accuracy and speed-confidence trade-offs are more striking for more complex tasks (Johnson et al., 1993; Hwang, 1994). Thus, I chose to expose participants to two tasks that differ in complexity: the relatively simple task of selecting the shortest route in distance, and the more complex reasoning task of selecting the fastest route in driving time. Time pressure and display types were both varied within the participants. However, a between-subject design was chosen for varying task difficulty, in order to avoid potential biases of learning and interference effects between the two different tasks. The research hypothesis in this context is: The effects of time pressure on response accuracy and response confidence are more striking for the more complex task, selecting the *fastest route*.

A further research question is how task type and display type might influence people's inference-making: On the one hand, the road network in the road maps should facilitate the task of solving the fastest route in driving time, because a classification of roads can give hints on possible speeds at which people can drive on certain roads. For instance, a motorway is usually represented by a darker color and a thicker stroke. As this road information is typically only available on road maps, one might hypothesize that people are more accurate with road maps than with satellite images. On the other hand, there is no reason to assume that participants would perform better with the road map when selecting the shortest route in distance, as the base map is irrelevant for measuring the distance of a path.

Another question in this context is whether people generally overestimate their response confidence when using realistic displays, as has been shown in previous work (e.g. Hegarty et al., 2009, see discussion in section 2.3.2). Furthermore, is response confidence dependent on sex? As prior work (Furnham, 2001; Furnham et al., 1999; Lloyd et al., 2002) suggests, males are more confident in their answers than females and tend to overestimate their response accuracy.

Finally, a detailed analysis of the route selection choices can shed light on the question whether the finding of Brunyé et al. (2010) that humans prefer southern over northern routes can be replicated for route selection choices under time pressure.

5.1 Participants

Seventy-six participants took part in this experiment, of which 44 (57.9%) were male and 32 (42.1%) female. The participants were mostly students and staff of the Department of Geography of the University of Zurich and the Institute of Cartography at the Swiss Federal Institute of Technology (ETH) Zurich.

Seventy-four of the 76 participants stated they would use maps occasionally (47.4%) or very frequently (50.0%) in their daily working life, while 70 also used them in their leisure time activities occasionally (78.9%) or very frequently (13.2%). It can therefore be assumed that my sample is familiar with using maps.

5.2 Materials

Participants were asked to select routes based on a randomized series of 24 map stimuli, which consisted of twelve road maps and twelve equally sized satellite images (400 x 400 pixels, see Figure 17) shown on a 17" desktop computer monitor. The stimuli represented twelve different flat urban environments from all over the world, in which selecting the shortest or fastest route from A to B was not a straightforward task. Figure 17 shows two map stimuli examples used in the experiment.

Each of the map stimuli consisted of a car symbol as origin (A), a person symbol as destination (B), three routes to choose from, and a base map representing urban environments from different locations all over the world. The roads were represented as polyline overlays and numbered with “1”, “2” and “3” in the corresponding colors.

The satellite images were horizontally and vertically rotated representations of the road maps, so that they represented the same geometrical problem and geographical setting as the road maps, but in a different orientation. Besides this rotation, the only other difference between the twelve pairs of images was therefore the base map - that is, the more realistic satellite image or the more abstract road map.

Figure 17: Test stimuli: road map (left) and rotated satellite image of the same area (right).

In terms of visual clutter (see section 2.3.3), the road maps are on average more cluttered than the satellite images. The average subband entropy for the satellite images is 3.20 (SD=0.19), while the average subband entropy for the road maps is 3.61 (SD=0.16). This difference is statistically significant ($p < .01$).

5.3 Procedure

Like Experiment I, this experiment was conducted via a web browser. It took place in a lab equipped with standard personal computers connected to the Internet, and was carried out with a standard web browser displayed in full-screen mode on a 17-inch color display set to 1280 x 768 pixel screen resolution.

After clicking a “Start” button at the welcome page of the experiment, participants were randomly directed to one of the two tasks, selecting the fastest road in driving time, or selecting the shortest road in Euclidean distance. 40 participants were directed to the shortest route task, and 36 to the fastest route task.

The experiment started with a pre-test questionnaire, in which users were asked to indicate their sex and map use experience in their private and professional lives. After filling out this background questionnaire, the experiment started with two warm-up questions that were identical to the task the participants would have to solve: selecting the shortest or fastest route from three choices on a satellite image or a road map. The choice for a road was made by clicking a radio button besides the options “route 1”, “route 2”, or “route 3” below the image.

After their route choices, participants were also asked how confident they were in their decision on a scale from 1 to 4, on which 1 indicated “not confident at all” and 4 “absolutely confident”. By using an even number of confidence levels, participants were forced to make a decision either with or without confidence.

Time pressure was simulated with a bar placed right of the map graphic, whose height was proportional to the seconds remaining, and whose color was either red (for 10 seconds), orange (for 20 seconds) or yellow (for 30 seconds), representing different levels of time pressure.

Each of these three time limits was assigned to eight of the 24 map stimuli for each participant. The order of the time limits was also randomized to prevent ordering biases. The experiment consisted of 24 selection tasks. Each of the three time limits (10, 20 or 30 seconds) was used 8 times, and each of the two display types (satellite image or road map) was used 12 times. With a screen resolution of 1280 x 768 pixels, as was available in the computer rooms in which the experiment was conducted, participants could see the map, the time limit and the checkbox for their decision at one glance on their screen without having to scroll. The overall duration of the experiment was approximately ten minutes.

5.4 Results

In this section, I will first discuss the effect of the context-related factors (time pressure and task complexity), then the effect of the map-related and user-related factors, and finally potential biases in road selection (northern vs. southern options).

5.4.1 Effect of time pressure and task complexity on response accuracy

Overall, response accuracy decreased with time pressure (30 s: M=44.2%, SD=26.2%, 20 s: M=42.1%, SD=23.5%, 10 s: M=38.6%, SD=23.2%, see “overall” columns in Figure 18). These results resemble a speed-accuracy trade-off. However, differences are altogether not statistically significant.

As Figure 18 shows, accuracy was higher when selecting the shortest route (M=55.6%, SD=14.4%) compared to the fastest route (M=26.2%, SD=8.8%), as expected. The accuracy differences between the two tasks are highly significant (p< .001). The low accuracy values of the fastest route task reflect that selecting the fastest route is obviously the more difficult task.

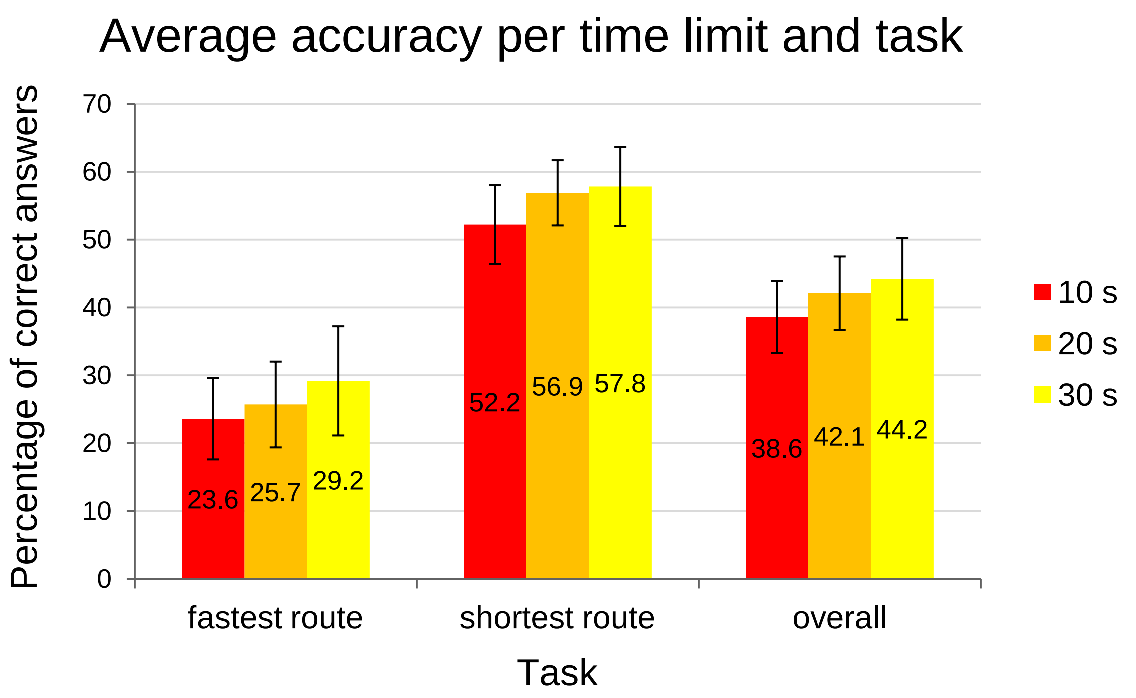


Figure 18: Average accuracy per time limit and task. Error Bars: ± 2 SE.

5.4.2 Effect of time pressure and task on confidence

Aggregating both tasks, an increase in response time limits seems to affect confidence in decisions (10 s: M=2.6, SD=0.5, 20 s: M=2.8, SD=0.4, 30 s: M=2.7, SD=0.5, see Figure 19) more than response accuracy. A repeated measures ANOVA indicates that the overall effect of time limits on confidence is statistically significant (p< .01). The same holds true for both the differences between the 10/20 s difference (p< .01) and 10/30 s difference (p< .05). However, there is no further confidence increase when participants have more than 20 seconds response time.

While overall confidence in responses is also higher for selecting the shortest route (M=2.7, SD=0.4) than for selecting the fastest route (M=2.6, SD=0.4), these confidence differences are statistically not significant (p> .05) (see Figure 19). There are significant interaction effects between time limit and task, regarding the confidence of results: On the one hand, there is an overall increase in confidence with a less severe time limit when participants performed the shortest route task (10 s: M=2.6, SD=0.5, 20 s: M=2.8, SD=0.4, 30 s: M=2.9, SD=0.4), which is significant (p< .01 for both differences in means between 10 s/20 s and 20 s/30 s time limits).

On the other hand, in the fastest road task there is no increase in confidence when people are given more time, and confidence even slightly decreases under the 30 second time limit.

A correlation analysis of overall accuracy and confidence values (Pearson's Rho = 0.22, p= .055) suggest that accuracy and confidence are not strongly correlated among participants. It can therefore not be assumed that people who make more accurate decisions also will be more confident in their decisions.

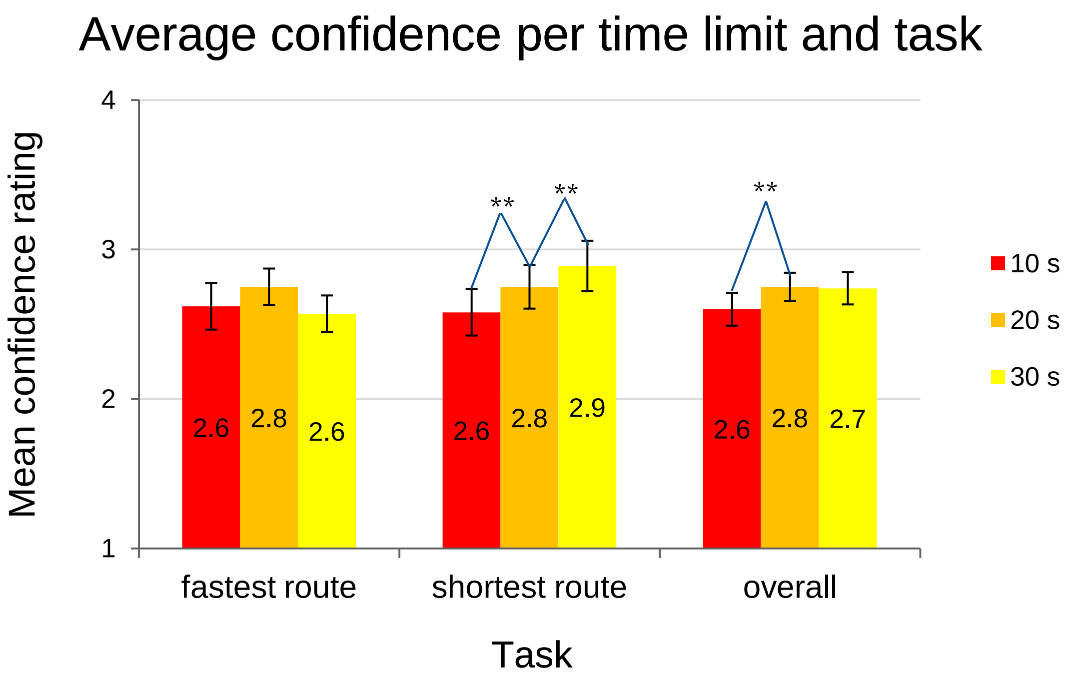


Figure 19: Average confidence per time limit and task. Error Bars: ± 2 SE, ** p< .01.

5.4.3 Effect of display type on response accuracy

Distinguishing between the two tasks "shortest route" and "fastest route", one can see that both road maps and satellite images have advantages for each of the two tasks (see Figure 20): Participants are more accurate with satellite images when selecting the shortest route (satellite images: M=58.5%, SD=17.6%, road maps: M=52.7%, SD=16.7%, p< . 05), while the participants who have to select the fastest route are more accurate with road maps than with satellite images (road maps: M=31.0%, SD=13.0%, satellite images: M=21.3%, SD=9.8%, p< .001), as expected. Aggregating the two tasks, accuracy values for both display types are similar and seem to balance each other (road maps: M=42.4%, SD=18.5%, satellite images: M=40.8%, SD=23.6%), and the differences between the two display types are statistically not significant (p> .05). The statistical significance of the interaction effect between task and display type, regarding the accuracy of results, can be shown with a repeated measures ANOVA (p< .05).

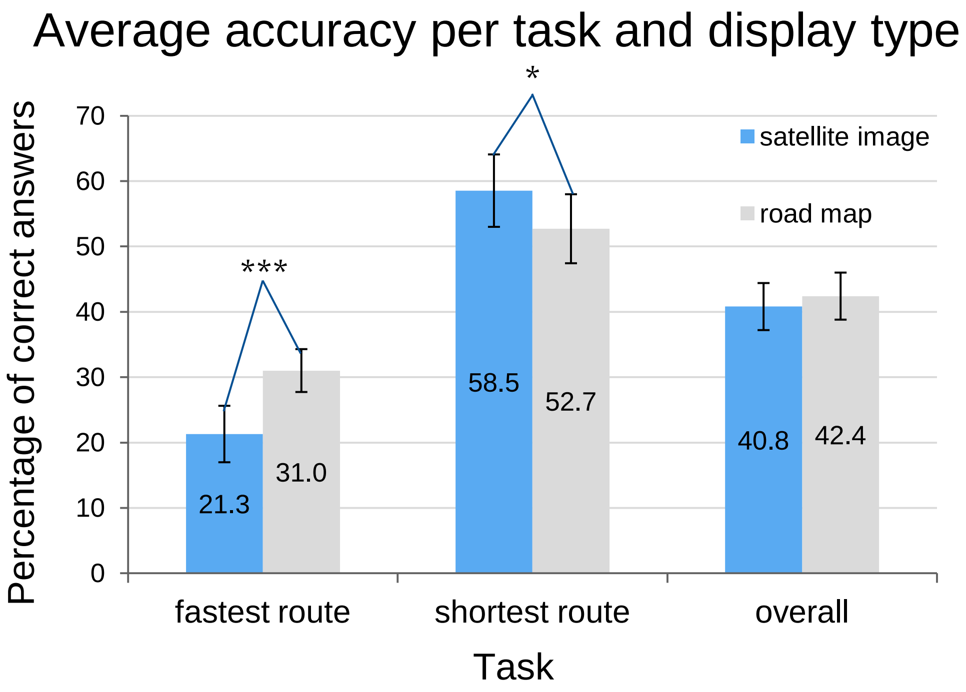


Figure 20: Accuracy per task and spatial display type. Error bars = ± 2 SE, *** p <0.001, ** p <0.05.

Another question in this experiment was whether there is a significant effect of time pressure on performance with satellite images, as the significant effects of time pressure on preference ratings from Experiment I would suggest. Aggregating the results from both tasks, one can see that the overall accuracy values for satellite images indeed are higher at the generous 30 s time limit (M=44.1%, SD=34.7%) compared to the two other response time limits (see also Figure 21). However, the average differences between the time limits are not significant for the satellite image. In contrast, for the road map, the overall accuracy increase from the 10 s (M=36.0%, SD=29.5%) to the 20 s (M=45.4%, SD=26.7%) time limit is significant (p< .05), while accuracy decreases again for the 30 s time limit (M=43.0%, SD=29.2%).

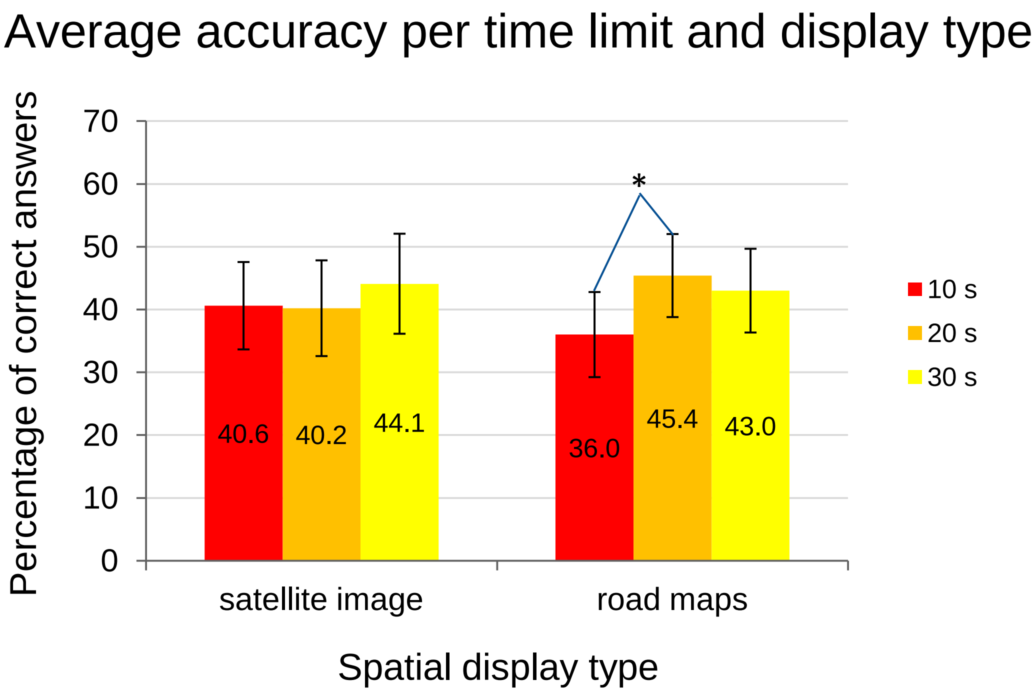


Figure 21: Average accuracy per time limit and display type. Error bars = ± 2 SE, * p< .05.

5.4.4 Effect of display type on confidence

In both tasks, shortest and fastest route, satellite images obtain higher confidence values than road maps (see Figure 22): In the shortest route task, mean confidence ratings for the satellite images are M=2.8 (SD=0.4), and M=2.7 (SD=0.4) for the road maps. Although participants are significantly less accurate in selecting the fastest route with satellite images, they are more confident in satellite images also for this task (satellite images: M=2.7, SD=0.4, road maps: 2.6, SD=0.4). Aggregating both tasks, average confidence in satellite images is M=2.7 (SD=0.4) and M=2.6 (SD=0.4) in road maps.

A repeated measures ANOVA indicates the statistical significance of the overall effect of display type on confidence ($p < .05$). These confidence differences between the map types are statistically significant for the shortest route task and overall ($p < .05$), but not for the fastest route task.

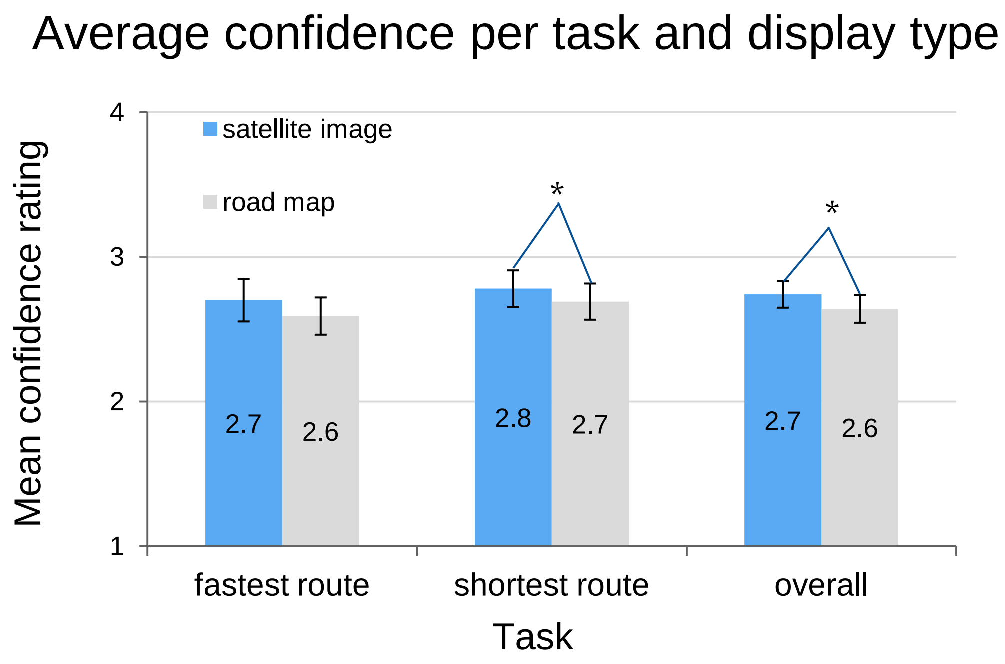


Figure 22: Confidence per task and spatial display type. Error bars = ± 2 SE, * p <0.05.

The confidence analysis per time limit and display types (see Figure 23) shows that confidence in satellite images is increasing from the 10 s (M=2.6, SD=0.6) to the 20 s (M=2.8, SD=0.5) limit (p< .01), but then also drops again to the 30 s (M=2.7, SD=0.5) limit. In contrast, confidence in the road map is generally increasing. Only for the 30 s time limit, confidence is higher in the road map than in the satellite image. These findings are not in accordance with the findings of Experiment I, where participants especially assigned higher suitability ratings to satellite images for the no time pressure condition.

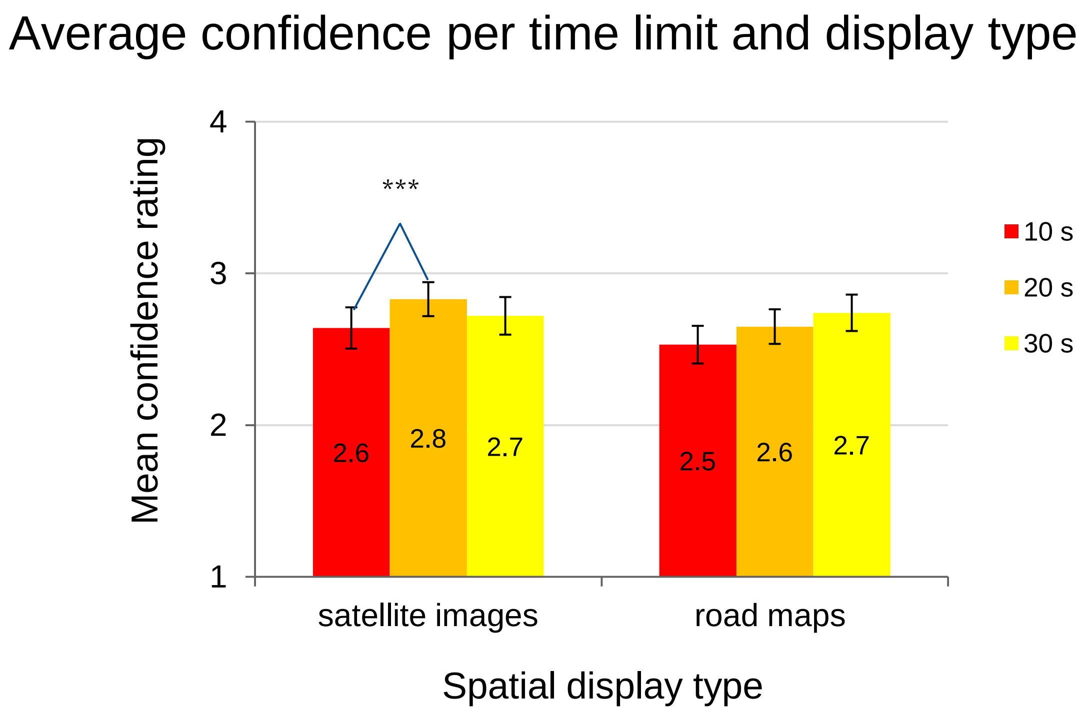


Figure 23: Average confidence per time limit and display type. Error bars = ± 2 SE, *** p< .001, * p< .05.

5.4.5 Effects of sex on response accuracy and confidence

As expected, sex has no effect on the accuracy of results. Females report higher accuracy values than males (female: M=43.2%, SD=18.3%, male: M=40.5%, SD=20.0%), but the differences are statistically not significant (p> .05). However, as hypothesized, male confidence values are higher (male: M=2.8, SD=0.4, female: M=2.6, SD=0.3), which is statistically significant (p< .05).

A repeated measures ANOVA reveals no significant overall interaction between sex and time. Taking a closer look at the single time limits, however, sex differences are significant for the short 10 s and 20 s time limits (10 s: male M=2.7, SD=0.5, female M=2.5, SD=0.4; 20 s: male M=2.9, SD=0.4, female M=2.6, SD=0.34), but not significant for the 30 s time limit. For female participants, confidence ratings generally increase with more response time, while male confidence seems to decrease when having more than 20 seconds for decision making (see Figure 24).

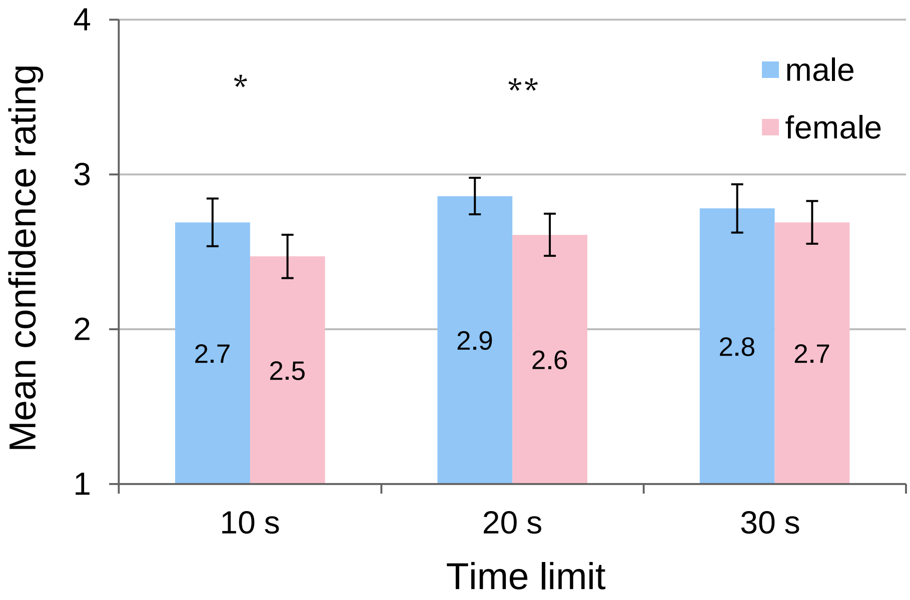


**Figure 24: Confidence ratings grouped by sex and time limits.
Error Bars: +- 2 SE, * p< .05, ** p <0.01.**

5.4.6 General effects of the independent variables (time, sex, display type and task type)

I conducted a repeated measures ANOVA to investigate whether the independent variables time, sex, display type and task type all have significant effects on the results, irrespective of the accuracy/confidence distinction. The analysis of main effects reports significant effects of task ($p< .001$), time pressure ($p< .05$), and sex ($p< .05$). The general effect of the independent variable "display type" is unclear ($p= .053$).

If the effects of the independent variables on the dependent variables accuracy and confidence are analyzed separately, quite a different picture can be seen: None of the four independent variables affect both confidence *and* accuracy. More precisely, the factors time, display type and sex all affect only confidence, but not accuracy. In contrast, the between-subject variable "task type" does not affect confidence, but only accuracy ($p< .05$). Altogether, three of the four independent variables affect confidence, but only one affects accuracy, and each of these four variables produce a certain significant effect.

5.4.7 Road selection biases

For twelve out of the 24 stimuli used, clear distinctions can be made between northern and southern route options, similar to previous work by Brunyé and colleagues (2010). As described in the *Materials* section (section 5.2), the original road maps and routes were vertically rotated in order to generate the satellite image stimuli. Thereby, correct Southern route options on the road maps became northern options for the satellite images and vice versa. Other properties, such as length or number of turns of the roads, were not changed (see Figure 25 for an example). Thus, I ensured that correct southern and northern routes were equally distributed. If participants chose southern routes more frequently, this would imply a preference for Southern routes, irrespective of accuracy.

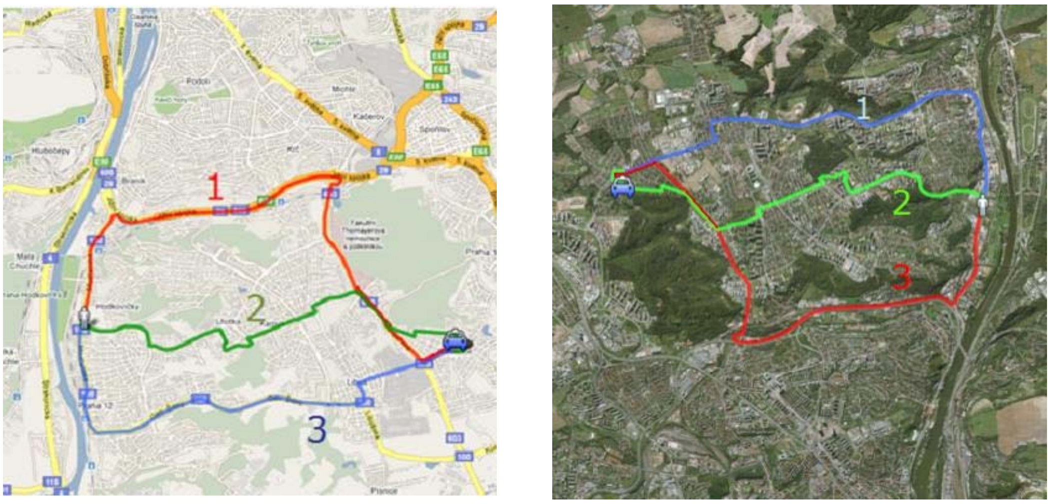


Figure 25: After rotating the original road map (left), the blue southern route option became a northern route in the satellite image (right).

However, the results do not suggest that participants would prefer southern over northern routes, but rather the opposite: On average, participants even more frequently chose northern routes (M=2.8, SD=2.2) than southern routes (M=1.8, SD=1.2). This tendency in favor of northern routes is significant ($p < .001$). Distinguishing between the two tasks, this north/south difference is only significant for the fastest route task, and not for the shortest route task.

5.5 Experiment II - Discussion

In this experiment, I measured people's accuracy and confidence under time pressure with satellite images and road maps. Accuracy and confidence was also compared with the preference ratings on the same task from Experiment I.

The results show that there is a slight increase in participants' accuracy with more time available, but this increase is not significant. Accuracy for a road selection task does not seem to be significantly affected by different time pressure scenarios (10, 20, and 30 second limit). This might indicate that the chosen decision time limits were not selective enough to influence task complexity (Johnson et al., 1993; Hwang, 1994), not even for the more complex task of selecting the fastest route.

However, the chosen time limits have an effect on people's confidence in road selection tasks, which is significantly higher when having more than 10 seconds to solve the task. The speed-confidence trade-off, found in research outside cartography (Maule, 1998; Maule and Andrade, 1997; Smith et al., 1982), seems to be more evident for this road selection task than the speed-accuracy trade-off. This, in turn, could imply that the accuracy with which participants select roads is actually less dependent on time pressure than they believe.

People's confidence when selecting the shortest route resembles a speed-confidence trade-off, while the results of the fastest route task are more similar to an inverted U-shape pattern. These results suggest that an increasing task complexity might lead to a U-shaped curve for confidence, while the less complex map-based decision making task follows more a linear speed-confidence trade-off pattern.

Aggregating the results for the two tasks, no overall accuracy differences between the tested road maps and satellite images for the route selection task can be detected. However, participants' overall decision making confidence is significantly higher when using satellite maps. One interpretation could be that participants tend to be overconfident in realistic displays (or underestimate their accuracy with road maps), as Hegarty et al. (2009) have found. This might be yet another indication of the *naïve realism* phenomenon (Smallman and St. John, 2005), i.e. novice users' misplaced faith in the utility of realism (see section 2.2). Another interpretation is that the labels in the road maps, such as Chinese letters, might have shown participants that they are indeed in an unfamiliar environment, and thus might explain low participant confidence in the road maps. In contrast, participants might have felt more confident in the satellite images, merely because the satellite images had no labels showing participants that they are in an unfamiliar environment.

Overall, accuracy and confidence in road selection tasks are not correlated, and there is no significant effect of sex on response accuracy. However, male participants tend to overestimate their confidence, as previous work has shown (Furnham, 2001; Furnham et al., 1999). This sex difference in confidence is especially striking under short limits.

Altogether, it could be shown that all four control variables, the two context-related variables (task and time pressure), the map-related variable display type and the user-related factor sex influence either the accuracy or the confidence in road selection choices.

Implications for follow-up experiments

One of the key findings of Experiment II is that time pressure seems to affect people's response accuracy more than their confidence in road selection tasks. An open question is how robust these findings about speed-accuracy and speed-confidence trade-offs are for tasks that are more complex than road selection in a flat urban environment. In a follow-experiment, it should be further investigated if the phenomena of overconfidence in realistic representations and the male overconfidence in accuracy can be replicated in another map use task with spatial displays varying in their degrees of realism. In order to increase task complexity and to test the performance with other map types from Experiment I - such as topographic maps or shaded relief maps -, the third dimension should be relevant for this map use task.

The next step within the framework of this thesis is to identify such a typical map-based decision making task for a follow up-experiment. This experiment should be similar to a real world context and shed further light on the research questions proposed in section 1.3. In order to understand expert decision making with maps and to learn which kinds of decisions and maps are relevant, I conducted interviews with experts in the field of map-based decision making, which are summarized in the next chapter.

6. Expert interviews: Map-based decision making under time pressure in the "real world"

After having measured map display and interactivity preferences for a road selection task (Experiment I, Chapter 4) and participant accuracy and confidence in road selection tasks (Experiment II, Chapter 5), the next step consisted of interviewing professionals, who perform map-based decisions under time pressure on a daily basis, specifically within a more complex three-dimensional context. These interviews can enhance the findings from Experiments I and II by putting them in a real-world-context, in order to inform an "ecologically valid" third experiment.

The focus of these interviews was on exploring what the concept of "time pressure" actually means for experts, how experts make map-based decisions under time pressure, and which spatio-temporal displays they use for their task. While experts might use certain map types for many years and thus might be familiar with them, other map types might be regarded as "nice-to-have" and desirable display types, and should be tested in further experiments.

The experts who were chosen and willing to participate in this study were ambulance drivers and dispatchers from Schutz & Rettung (Protection & Rescue) Zürich, mountain guides from SLF (Swiss Institute for Snow and Avalanche Research), urban police officers (Stadt Zürich), employees of the electric power company of Zurich (EWZ), and finally helicopter pilots from Swiss Air-Rescue (Rega). The interviews were conducted in March and April of 2010.

Military is another important area where map-based decisions under time pressure have to be made frequently (Torun and Ulubay, 2005). Although the scope of this thesis did not allow for in-depth interviews with military personnel, I will discuss a source which sheds light on the use of maps under time pressure for military purposes in the final section of this chapter (section 6.8).

6.1 Dispatchers (Schutz & Rettung I)

Schutz & Rettung ("Protection & Rescue") is the largest rescue organization in Switzerland. Its areas of operation include fire service, medical service, operation control centers, fire prevention police and civil defense (Schutz & Rettung Zürich, 2010a). It is the umbrella organization for several departments, such as rescue service, fire brigade, civil defense, operations control center and fire police, and it is integrated into the emergency response organization of the city of Zurich, the canton of Zurich and the Swiss Federation. Its mission is to provide the best possible and the highest level of protection for the population in the city and metropolitan area of Zurich.

This mission also involves operating under severe time pressure, as illustrated by a statement of the brochure: *"People's lives are dependent upon the earliest possible emergency medical treatment".* (Schutz & Rettung Zürich, 2010b).

My interview partner was the deputy department head of the Operation Control Center (OCC) in Zurich-Wiedikon. The rationale for choosing her as an interview partner was that her everyday task is very similar to the scenarios of my previous experiments: ensuring that "the appropriate assistance reaches the right place in the shortest possible time". The instruction for Schutz & Rettung is that the ambulance drivers are at the place of action within 15 minutes. In 2007, the medical service performed approximately 90 tasks per day.

Map use routine

On a typical shift at the OCC, six dispatchers receive and handle emergency calls. Each dispatcher sits in front of four computer monitors displaying the database and map components of a geographic information system. In the unlikely event of power failure or of a large-scale disaster that would hinder access to the GIS, the dispatchers also have large-scale topographic paper maps and schematic plans of important buildings such as the Zurich central train station at their disposal in the OCC.

Most of the dispatchers are paramedics, who are not specifically trained in cartography. Map-reading is regarded as a prerequisite when starting the job. Being flexible, having short reaction times, staying calm in hectic situations, and having good communication skills are seen as more important than map-reading skills. The dispatchers' work, which consists of immediately finding the place of action based on telephone calls from citizens, is often carried out under severe time pressure. The statement of one of the dispatchers that even a search in Google Earth would take too long for this task supports this view.

The working routine in the center is as follows: The dispatcher receives a phone call and has to assess to which location (which in most cases is the location of the caller) the emergency vehicle has to drive. After the dispatcher has determined this location and entered it into the GIS database, the map display is centered at this location. As a consequence, for the dispatchers only one question is relevant: "Where?" In other words, their task is georeferencing the place of action. This can be difficult in the case of imprecise location information (Lang, 2010). One example for an imprecise location is the statement *"The accident happened* *somewhere on the motorway between Sankt Gallen and Basel",* in which the possible place of action is somewhere within a motorway section of 166 km length.[2]

Topographic maps from swisstopo are used to query the database of the GIS, which contains the two typical interaction tools for a 2D map, panning and zooming. Depending on the zoom level,, the level of detail on the map changes. However, contour lines with elevation information are visible at all map scales. Only at a certain zoom level, road names and house numbers are visible. Within the urban areas, the dispatchers can change the display type to satellite images. However, according to one of the dispatchers, they almost never use the satellite images.

My interview partner also mentioned that the third dimension and 3D displays are irrelevant for their work. She mentioned she would indeed use Google Earth and Google Street View[3] (which had just been released for Switzerland at the time of the interview) in some cases for her work, but regarded the loading times (display refresh rates) as too slow. However, she supposed that the third dimension would matter for helicopter pilots (see section 6.5).

The database part of the GIS consists of *TeleAtlas* road data, and the database query was initially only working with addresses. However, the staff of Schutz & Rettung have also integrated locations without addresses into their database. In other words, they have address-matched places that do not have a street address but frequently occur in emergency response scenarios (such as swimming pools).

After the dispatchers have tried to answer the "Where?" question and located the place of action, they hand over the map and the geographic coordinates to the paramedics (see section 6.2). The coordinates are mainly used for the in-car navigation device, whereas the map serves as a static graphic display, always printed at a fixed scale (1:20,000). A scan of a sample map which has been handed over from the dispatchers to the paramedics is shown in Figure 26.

[2] As calculated on http://maps.google.com
[3] http://maps.google.com/help/maps/streetview/

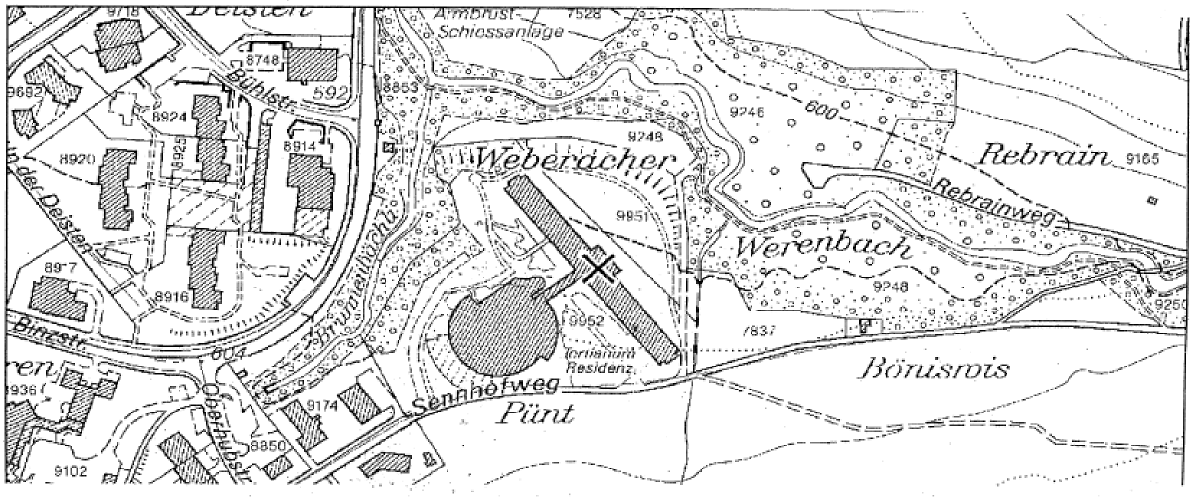


Figure 26: Sample map used by the ambulance driver.

Map use requirements

When asked which information maps should contain for their task, the dispatchers mainly mentioned road data and road names. It is important that the map is well designed, easy to read, and that buildings are easy to detect. On the whole, the dispatchers are satisfied with the map material they have at hand. Additional 3D information was regarded as “nice to have”, but none of the four interviewed dispatchers were of the opinion that virtual globes would substantially facilitate their work. However, one of them could think of a decision making scenario where a virtual globe might facilitate their working routine. This scenario consists of deciding whether an all-terrain vehicle should be employed in a rescue operation in a mountainous area.

The dispatchers were also asked to rate their map display preferences regarding the six maps used in Experiment I. They agreed that the two road maps (Map C and E) were most suitable, whereas they regarded the topographic map (Map A) and the shaded relief map (Map B) as least suitable for their task of detecting the place of action. This is congruent with the results of my Experiment I (Chapter 4) with novices in search & rescue.

6.2 Paramedics (Schutz & Rettung II)

Following the interviews with the dispatchers, I interrogated a shift of six paramedics of Schutz & Rettung, in order to assess in which aspects their everyday usage of maps differs from their colleagues in the OCC. A first interesting statement of the ambulance drivers was that they only rarely use the sketch map (see Figure 26) they receive from the dispatchers. However, this sketch map contains information about the respective grid square in the large-scale printed paper city map. This city map is still the most commonly used map in the field - even in times where GPS technologies and maps on mobile devices seem to have become ubiquitous and increasingly popular.

The ambulance vehicle itself contains a standard navigation device. The paramedics mentioned that this navigation device often suggests “nonsense” routes. Thus, they prefer to rely on their own route, landmark and survey knowledge (Goldin and Thorndyke, 1981) of the area in most cases. Moreover, they regard their GPS as too slow.

About two minutes pass on average between receiving the order from the OCC and the ambulance drivers being on a route to a target destination. According to one of the paramedics, the decision about which road to take has to be made within 40 seconds. There is no instruction when to be at the target location (*“this is my business”),* as it is obvious that driving as fast as possible also bears some risks of additional accidents.

The paramedics share the dispatchers’ opinion that the road maps (Map C and E in Experiment I, see Figure 7) are the most useful ones for their purpose. Satellite images are regarded as helpful in areas without road names, like forested hills (such as the Uetliberg). The paramedics also regard the third dimension as irrelevant for their task: All six paramedics interviewed agree that they never pay attention to the contour lines on the maps – which are featured in all maps, as mentioned before.

6.3 Local police (Stadtpolizei Zürich)

Map-based decision making under time pressure is also a common task for urban police departments. For instance, police officers might have to reach a site of crime as quickly as possible. My interview partner in Zurich was the head of the operation control center of the city police. In this center in Zurich-Wiedikon, the city police use the same Intergraph GIS as Schutz & Rettung. However, my interview partner claims that the use of the GIS system is more complex than for Schutz & Rettung, because the police dispatchers do not only have to determine a target location, but also the needed gear that the police car drivers take to the place of action.

Apart from that, working with the map under time pressure is very similar to the task of the dispatchers at Schutz & Rettung and thus requires no detailed further description.

6.4 Mountain guide (Institute for Swiss Avalanche Research, SLF)

My next interview partner works as a mountain guide for the Institute for Snow and Avalanche Research (SLF) in Davos in the Swiss Alps. He is a geographer by training, spends a lot of his working time in the field collecting data and analyzing accidents, and also develops GIS-based prediction systems, such as an interactive 3D avalanche training tool.

According to him, the classical topographic paper maps are still the prevalent orientation device for mountain guides in the field. While mobile devices with digital topographic maps are used frequently, the fact that these displays are hard to see under snowy or sunny conditions complicates their usage. Moreover, most mountain guides do not like to be dependent on electronic devices. For instance, the power supply must be guaranteed. This is not an issue with paper maps, which are therefore more reliable in this respect.

Regarding their requirements for maps for mountaineering purposes, mountain guides are quite a heterogeneous group. Most of the guides still prefer classic topographic paper maps, and some of them do not use any GIS at all. However, some mountain guides regularly load GPS tracks on their mobile phones and use them in order to be more flexible when selecting routes. However, these cases can be regarded as exceptions. Planning a mountaineering tour requires a "serious" preparation, and it is highly doubtful whether this is possible with using skiing tour GPS tracks from websites whose content is based on VGI (volunteered geographic information), and about which not much metadata are known. This heterogeneity of map use preferences could be due to the fact that the background of the mountain guides varies a great deal. As was the case for the dispatchers, map reading is an ability which is regarded as mandatory when starting the job of a mountain guide.

As for the drawbacks of topographic maps, the mountain guide firstly mentioned the depiction of forest areas, because the topographic map does not give information about the density of the trees, and secondly, the occurrence of crevasses, which are mostly not adequately depicted on topographic maps. For this purpose, a thematic map with explicit slope information would be the perfect map type. As slope is the critical factor for predicting avalanches (Suter, 2007), a specific slope map could give additional insights on avalanche danger. With this map display type, it could be ruled out that mountain guides cross an area with a critical slope value. An example of such a map is the Ski Tour map from swisstopo[4], on which slopes over 30 degrees are shaded in red colour.

A mountain guide typically does not work under as severe time pressure as a dispatcher or ambulance driver (see sections 6.1 and 6.2). In most cases, his routes are planned one day in advance on the computer screen or a paper map. However, there are occasions when mountaineering groups are travelling at high velocities, such as ski-tour free riding. In these

[4] http://www.swisstopo.admin.ch/internet/swisstopo/en/home/products/maps/leisure/ski.html

cases, a map-based decision has to be made within one or two minutes. For this purpose, a 3D model, which visualizes the steepness of the terrain and thereby gives quick advice on whether and how much the path is going uphill or downhill, would be more suitable than a paper map. In addition, the third dimension matters also for avalanche dynamics, as mentioned in the previous paragraph.

When asked whether Google Earth with its explicit depiction of the third dimension was a publicly available alternative to current state-of-the-art maps used for decision making, the mountain guide mentioned that the satellite images in Google Earth were too imprecise for his needs. Only if the images in this program were at a high resolution, would they be of any use for a mountain guide. He doubts whether mountain guides generally would prefer working with a photorealistic 3D model over a topographic map. In addition, in bad weather conditions or when snow is covering the area of action, the satellite image might not give a realistic depiction of the current situation in the area.

Looking into the future, the mountain guide could envision a technology, by means of which the combination of a person's current GPS position with a 3D model would generate a visualization of this person's view.

6.5 Electric power company (EWZ)

Next, I interviewed two engineers at the electric power company *EWZ*, which is in charge of supplying the city of Zurich and parts of the canton Grisons with energy since 1892 (EWZ, 2010). EWZ was chosen as an interview partner, because they also make map-based decisions in their every life, and have to be prepared for time pressure scenarios such as a total urban electricity blackout involving collapse of traffic lights and trams, which would make a fast restructuring of the entire electricity network necessary.

EWZ runs an Operations Control Center which is similar to Schutz & Rettung. Their dispatchers use "pseudogeographic" schematic network displays with topologic maps of the electricity supply network of Zurich. The most important transformer substations and power supply lines are depicted on these maps. In addition, a GIS by Smallworld is in use. If desired, a topographic map and an official city map can be added to the visualization of the electricity network. Since 2000, this GIS has replaced paper maps (at a scale of 1:25,000), which are however still used as a backup.

EWZ chiefly uses maps to locate blackouts. After localizing a blackout, an engineer drives to its location and tries to repair it. The engineer has to be at a target location within four hours, so the time pressure is not as severe for EWZ as for Schutz & Rettung. However, in the worst-case scenario described at the beginning of this section, the time pressure for making effective and efficient map-based decisions would be very severe.

When asked about map use preferences for their tasks and the relevance of the third dimension, our interview partners mention that the third dimension is only important for EWZ in order to assess how deep the electric cables are located below the Earth's surface. Realistic images, such as satellite images, are regarded as "nice to have", but what matters for EWZ is finding out the location of the electric cables and the houses that are supplied by the electricity from these cables. As for the maps used in Experiment I, the dispatchers mentioned they would prefer the road maps because they are familiar with them, while a 3D map might be useful in mountainous areas of the supply network, such as the Grisons.

6.6 Helicopter pilots (Rega I)

My final interview location was at the headquarter of *Rega*, also called "Swiss-Air Rescue". Rega is an independent, non-profit-foundation and member of the Swiss Red Cross (Rega, 2011). Its primary mission is to provide medical assistance to the scene of an accident by helicopter, and then a secondary mission is to transfer patients to the hospital. The organization maintains twelve bases in the area of Switzerland. Its network is structured in such a way that an emergency helicopter (see Figure 27 for an example) can reach the place of action within 15 minutes. As Rega states it on its website, *"every minute could mean the difference between life and death."* [5] The helicopters fly at a speed of up to 270 km/h, and Rega conducts about 10,000 rescue operations per year. Rega is mostly operating in mountain areas, but also in urban areas.

Figure 27: Rega helicopter (Source: Rega [6])

[5] http://www.rega.ch/en/rega/start_rega_einsatzschweiz.aspx?pid=040203000000
[6] http://www.rega.ch/en/about-us.aspx

My first Rega interview partner at Zurich-Kloten airport is the head of the paramedics department. He has been working at this organization since 1987.

Map use routine for the helicopter pilots

The helicopter pilots have to enter the geographic coordinates they receive from the dispatchers (see chapter 6.7) manually into the GIS. Five minutes after they have received the order from the Operations Center, the pilots have to be ready for departure.

The current cartographic state-of-the art for the helicopter pilots at Rega is a set of so-called *Air Navigation Obstacles Maps*[7]. The base map of this map type is a topographic map at the scale of 1:100,000, to which an additional layer of air navigation obstacles, helicopter landing fields, air fields and hospitals is added. This layer includes all relevant barriers and perils for the pilot, such as cable cars, cables, towers, electricity stations, which are at least 25 meters above the ground. This paper map is the main cartographic source at Rega. According to the head of the department, one is just more flexible with a paper map, in case of the collapse of an electronic device or bad weather conditions.

This Air Navigation Obstacle map is also used digitally in the helicopter cockpits on a navigation device, whose display is about 15 x 20 cm in size. The current position, elevation and flying route are also transmitted to the navigation device. On a smaller screen, an instance of *EURONAV V*[8] with a (non-interactive) 3D visualization and *Moving Terrain*[9] is running.

At the moment, it is not possible yet for Rega to rotate their cockpit maps in direction of travelling; they are always north-up. The Swiss military, however, are already using overhead displays and simulate a flight through a virtual valley with a 3D view.

When flying at night or in other poor visibility conditions during the day, Rega pilots additionally use heat images from infrared cameras and display them beside their obstacle map. In addition to these heat maps, they receive acoustical signals when approaching potential obstacles.

Map use requirements and relevance of the 3rd dimension

My interview partner regards an interactive 3D visualization as basically desirable. However, like the mountain guide (cf. section 6.4), he states that the usage of such 3D maps would only make sense if the resolution of the satellite images was higher than it currently is, and if the

[7] http://www.swisstopo.admin.ch/internet/swisstopo/en/home/products/maps/aero/obstacles.html
[8] http://www.armaviation.com/index.php?page=euronav-v
[9] http://www.moving-terrain.de/

servers rendering the images would be faster than, for instance, Google Earth currently is. Moreover, he regards the shadows on the images often as disturbing. He predicted that in five to ten years 3D models might be the state of the art, but not now.

While the working routine at Rega is similar to Schutz & Rettung in some aspects, the main difference lies in the demand for map displays: For Rega, the third dimension plays a crucial role. This does not only concern the flying route, but also the slope of the terrain: If the slope is too steep, the helicopter cannot land, and a rope winch needs to be used. The head of department at Rega agrees with the mountain guide that slope maps would substantially help to facilitate the task of his team, and provide a benefit compared to the current maps. Apart from that, he is quite satisfied with the currently used maps. Additional content besides slope information would not be helpful in his opinion. His pilots are used to topographic maps, and "like with everything", a switch from 2D maps to 3D maps would take its time. This statement is in accordance with my findings from Experiment I, where participants mentioned that they preferred familiar visualizations under time pressure.

6.7 Helicopter dispatchers (Rega II)

After speaking to the helicopter pilots, I was also given the opportunity to visit the Operations Centre and speak to one of the dispatchers.

Map use routine

The working routine for the dispatchers at Rega is very similar to the one at Schutz & Rettung: The Rega Operations Center at Zurich Airport receives phone calls, and the dispatchers detect the target location on the map via using the interaction tools zooming and panning. Then, they forward the map to their colleagues, who try to reach the target direction as quickly as possible. However, unlike at Schutz & Rettung, address data are only of minor importance at Rega. As Rega mostly operates in mountain areas, exact geographic coordinates are required. Therefore, locating the place of action is more difficult for helicopter pilots than for an ambulance mission. At the time of the interview, Rega did not use any GPS positioning technology for locating calls via mobile phones yet.

The Rega dispatchers use a different GIS than the one used at Schutz & Rettung, EWZ and the city police. The geodatabase part consists of data by SwissNames[10] and important data georeferenced by Rega themselves, such as mountain shelters or restaurants. Aerial obstacles, as displayed in the hardcopy Air Navigation Obstacle Maps, are not visible on the dispatchers'

[10] http://www.swisstopo.admin.ch/internet/swisstopo/de/home/products/landscape/toponymy.html

displays, as visualizing these line objects would technically not be possible. However, the map includes points of interest (POI), such as gas stations, mountain shelters, hospitals and helicopter landing spots, as well as scanned tourist maps of skiing areas. The dispatcher mentions a certain learning effect in a sense that he now finds places more easily now than back in the days when he started his job.

Unlike at Schutz & Rettung, Rega does not own satellite images, for the costs of acquisition of satellite data (10 million Swiss francs) is regarded as too high, and, when used in an interactive GIS, these images are also regarded as too little performant for their dispatching tasks under severe time pressure.

Role of time pressure

According to one of the dispatchers, he and his colleagues are under a more severe time pressure than the Rega pilots. In most cases, the time between receiving the phone call and transmitting the geographic coordinates to the pilot is between 30 seconds and 10 minutes, when the information provided by the caller is imprecise and fuzzy. For instance, in one typical case locating a place called "Drümännler" in the Berner Oberland region took ten minutes, because for georeferencing via the database the exact notation of the name is necessary, which is often not straightforward in Switzerland with its huge variety of dialects and notations of places. Due to this fact, a good geographical knowledge of Switzerland is necessary for their jobs.

Map use requirements and relevance of the 3rd dimension

Like the pilots, the dispatcher also regards 3D visualization as "interesting" in principle, but sees no adequate way of using it that would justify paying the high costs for Rega at the moment. While he indeed uses Google Earth for his work sometimes, he mentions that he cannot rely on the accuracy of the information. In addition, he refers to an aspect which has not been brought up yet, namely, that it is desirable if the caller and the dispatcher use the same cartographic material. For instance, if the dispatcher communicates with a mountain guide, who can determine that his current location is "the S of Druesberg" on the 1:25,000 map, it is rather easy to detect the exact coordinates. This is another argument for the preference for and usage of familiar maps under time pressure.

As for the maps used in Experiment I, the dispatcher mentioned that the shaded relief map would be of little use for Rega purposes. While the 3D map (Map D) would be "nice to have" and helpful for getting a first overview of the area, it has to be at a high resolution in order to be really useful, which also holds true for the satellite image (Map F). The generally preferred road maps C and E would be useful for establishing the region of the accident, as they include place names, which are easily readable. These are important issues of map design which have to be taken into account when assessing the effectiveness and efficiency of maps.

6.8 Military (German Armed Forces)

Efficient map-based decision making under time pressure is crucial also for military purposes. Dahlke and Winck (2009) give an overview about maps currently in use at the German Armed Forces (*Bundeswehr*). These authors make the claim for a solution similar to Google Earth or NASA World Wind[11] for military purposes and show how this is done in-house at the Bundeswehr Geoinformation Service. They also give examples (see Figure 28) which support the view held by my interview partners that the map data in Google Earth are not at a sufficient resolution for their task. In areas where military is typically operating (they mention Kosovo, one could also think of Iraq and Afghanistan), the poor coverage with publicly available high-resolution satellite images is even a more severe problem than in areas such as Switzerland or the United States.

[11] http://worldwind.arc.nasa.gov/java/

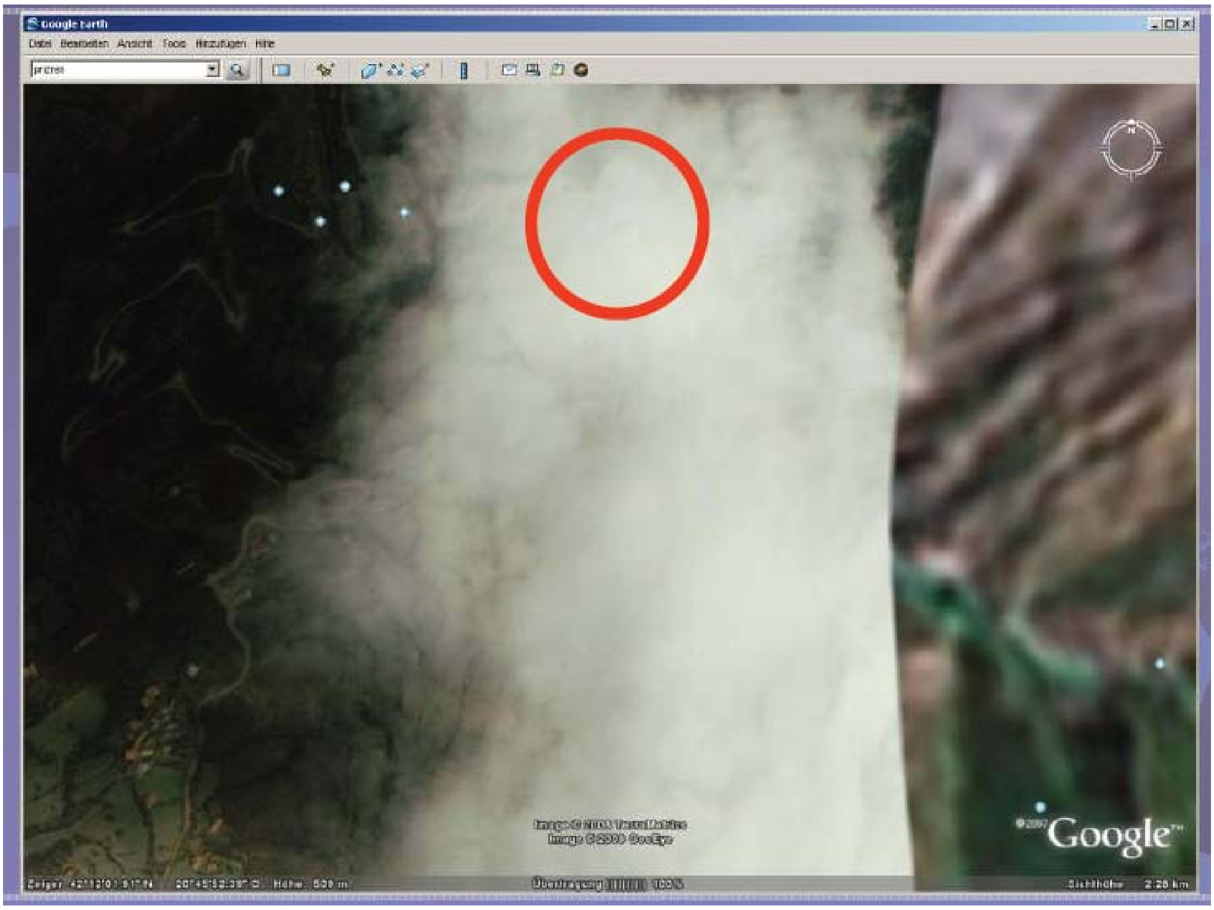


Figure 28: A region in Kosovo which is relevant for a military operation (Dahlke and Winck, 2009). Due to the poor visibility of relevant objects, this view obtained in Google Earth is not useful for map-based decision making under time pressure.

Therefore, the German Armed Forces acquire their own data, on the basis of which they generate so-called image city maps (ICM) for a quick first overview of the operation area, and finally their own visualizations by adding several other layers. Figure 29 shows the result of such an in-house visualization at a very high resolution, which supports the statements of the mountain guide and the Rega dispatcher that satellite images have to be at a sufficient resolution when used for map-based decision making under time pressure.

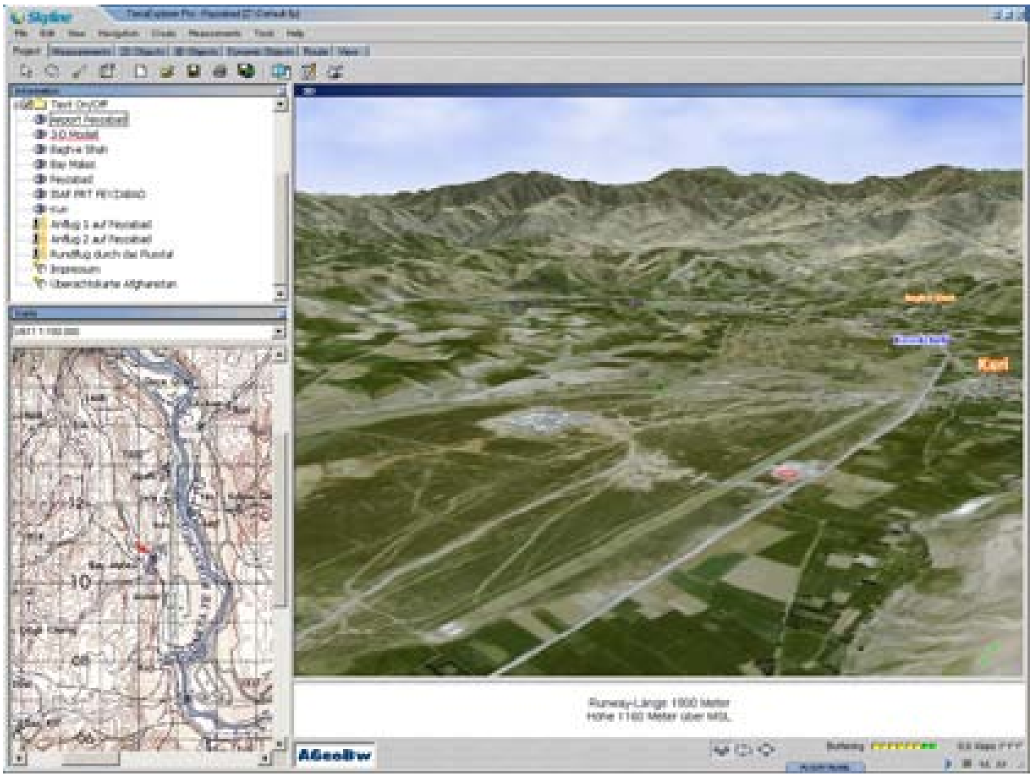


Figure 29: 3D simulation of an airport in Afghanistan (source: Amt für Geoinformationswesen der Bundeswehr[12]). The objects are at a resolution which is high enough for military purposes, in contrast to the poor resolution in Figure 28.

6.9 Summary

My expert interviews provide counter-arguments to the widespread belief that 2D static paper maps are outdated in times of location-aware mobile displays and 3D virtual globes. The majority of experts stated that none of these innovations can currently fully replace a large-sized classic static 2D topographic paper map for several reasons, which I will further discuss in the following paragraph.

Firstly, under time pressure there seems to be a preference for familiar displays, which supports the results of Experiment I. The most familiar map type for expert decision-makers under time pressure is (still) the topographic map. The majority of search and rescue personnel have been specifically trained with these maps, can read them well, and thus are generally comfortable with using them. Secondly, there are severe display limitations of mobile devices – one can simply see more on a large paper map, and visibility is limited under certain viewing conditions, such as plain daylight. Thirdly, publicly available satellite images on digital displays (e.g., in globe viewers) are not at a sufficient resolution. Using the categories introduced in chapter 1.2, the image resolution can be regarded as a "within-spatial-display-

[12] http://www.streitkraeftebasis.de/portal/a/streitkraeftebasis/dienst/portraits/ageobw

type" issue of map design, and demonstrates the influence of map design on map-based decision making. In many cases (i.e., Rega), the advantages of acquiring expensive, high resolution satellite images, do not compensate for the high costs. In other cases (i.e., military), high-resolution satellite images are produced in-house. Finally, expert decision-makers tend to be reluctant to rely on a display which might be unusable or not accessible at all times, i.e. due to power failure. Finally, these digital maps are at the time of writing often not fast enough.

A "slope map", that is, a thematic map which provides explicit slope information, is regarded as useful (but unusual) by expert map-based decision-makers under time pressure. When flying a helicopter or doing mountaineering, accurate slope identification is very important. For example, a helicopter must assess the steepness of the terrain for landing, while for a mountain guide the steepness of a slope needs to be regularly assessed for determining the avalanche potential when on a ski tour during the snow season (see also Suter, 2007). Although most of the experts mentioned that they were generally satisfied with the topographic maps used in their daily professional lives, mountaineers and helicopter pilots regarded a slope map as "nice to have", and see future potential for this still unusual visualization method.

Implications for further work

While the expert interviews have confirmed that route selection is a typical map use task under time pressure, slope detection emerges as another typical, more complex task, which should be tested in an experimental setups. As most of the experts have mentioned that topographic maps are the state-of-the art for map-based decision making, it should also be tested whether participants actually perform better with topographic maps than with maps in which the thematically relevant third dimension is depicted differently, such as the maps tested in Experiment I (shaded relief maps) or maps which experts would regard as particularly useful for their tasks (slope maps).

As some experts also mention that not only static, but also interactive GIS are used in their working routines, an open question is how the efficiency and effectiveness of decision making under time pressure depends on interactivity. This research question will be explored in Experiment IV. Prior to that, Experiment III will focus on a slope detection experiment with static maps.

7. Experiment III: Slope detection with 2D contour and 3D relief maps

In Experiments I and II, a significant effect of time pressure on user preferences and response confidence for realistic 3D-looking maps in a 2D task context could be uncovered. However, actual performance did not seem to be affected by the verisimilitude of the display to the real world.

Experiment II (Chapter 5) focused on a decision making context (road selection in a 2D environment), for which the third dimension was irrelevant. In Experiment III, reported in this chapter, I specifically investigate decision making within a 3D context and different display types, and analyze collected responses (i.e., accuracy) within the signal detection theory framework. The main research question for this experiment was how 3D realism might affect participants' response accuracy and confidence for a task under time pressure that specifically involves decision making within a 3D context. Since the expert interviews with helicopter pilots have shown that slope detection is a typical map-based decision making task under time pressure, this task was chosen for Experiment III. As the task of slope detection (in a 3D environment) can be regarded as more complex than route selection (in a 2D environment), the main assumption is that the speed-accuracy and the speed-confidence trade-off might also be more significant in this slope detection task than in the road selection task.

The exact scenario for this experiment was to identify locations on maps where a helicopter can land. Experts at Rega mentioned critical slope values of 14% or 8 degrees for safe helicopter landing, which is a value also stated in other sources (Bloom, 2007).

7.1 Participants

Fifty-five (32 male, 23 female) participants took part in this experiment. Like in Experiment II, participants mostly consisted of staff and students at the Department of Geography at the University of Zurich, and the Department of Cartography at the Swiss Federal Institute of Technology (ETH) Zurich.

The majority of the participants stated they were "rather familiar" with topographic maps (58.2%) and 3D maps (61.8%), while 32.7% reported they were "very familiar" with topographic maps, and 14.5% specified to be "very familiar" with 3D depictions. Therefore, one can assume that the sample is typical of the experienced map user under time pressure and similar to Experiment II. While this sample represents the more experienced map designer and user, the participants are not experts in detecting slopes, and do not represent experts in map-based decision making under time pressure either.

7.2 Materials

Twelve map stimuli were created for this experiment. Each stimulus had the identical map size of 389 x 355 pixels plus a reference scale at the right of the map and showed mountainous areas in Switzerland. The elevation data for the stimuli were derived from the SRTM3 Digital Elevation Model (Jarvis et al., 2008). With these maps, slope could be calculated from two pieces of information: the scale of the map and the contour lines. The scale of the map was held constant at 1:20,000 (run), while the interval of the contour lines was held constant at 100 m (rise). Our twelve map stimuli (see Figure 30) can be divided into four different classes of topographic maps with the following slope representation:

1. Only contour lines (Map A)
2. Contour lines plus light shaded relief (Map B)
3. Contour lines plus dark shaded relief (Map C)
4. Contour lines plus colored slope map (Map D)

While all map types are suitable for identifying slope, the depiction methods systematically vary in the degree of the depicted realism (i.e., shaded relief vs. contour line maps), in the apparent visual clutter (i.e., contour line vs. shaded relief maps), and in the information content for detecting slope suitability (slope vs. contour line maps). In other words, the maps are neither computationally nor informationally equivalent (Fabrikant et al., 2008; Simon and Larkin, 1987). The map with only contour lines represents a frequently used topographic map, while the two relief maps contain relative-relief portrayal methods as additional pieces of information. These methods are quite common in topographic maps for mountainous areas (Collier et al., 2003; Kimerling et al., 2005).

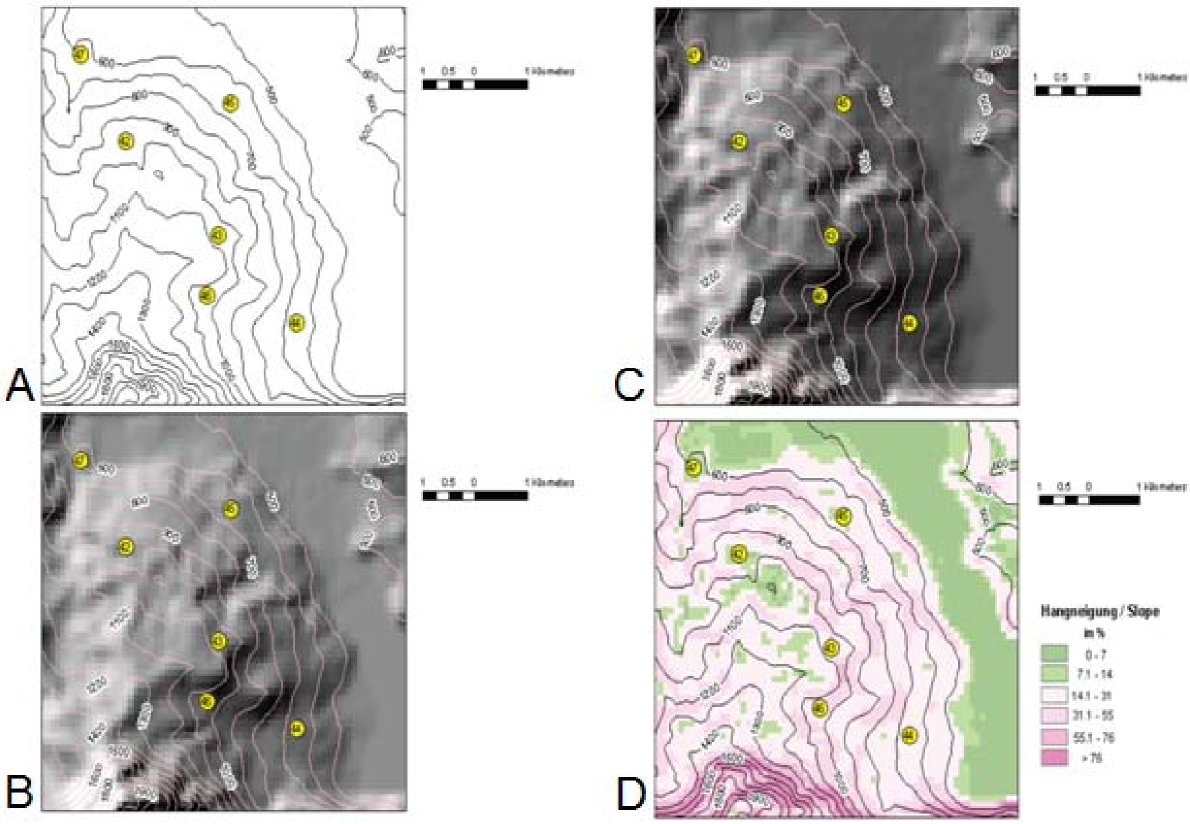


Figure 30: Reduced examples of map stimuli used: only contour lines (A), light shaded relief (B), dark shaded relief (C), slope map (D).

Map A in Figure 30, containing only contour lines, represents the most abstract of the tested map types, and with the least amount of information (i.e., implicit slope information). Maps B and C additionally contain a shaded relief (i.e., explicit relative slope information), thus more information than Map A. Users can obtain slope information not only (implicitly) from the distance between the contour lines (Map A), but also from the relative darkness of the pixels (Maps B and C). The appearance of the relief is influenced by the azimuth (i.e., horizontal angle) and the vertical angle of the light source. To investigate the potential effect of the brightness of the hill shading, I created a lighter version (Map B) and a darker version (Map C) of the hill shaded relief maps. I employed the hillshading function available with the 3D Analyst Toolbox in ESRI's ArcGIS. The light source for the hillshading was set to a 45° vertical angle for the lighter relief (Map B) and to a 22° vertical angle for the dark relief (Map C), respectively, without changing the azimuth. Based on Simon and Larkin (1987), I hypothesize that the more implicit the depiction of the task relevant information (i.e., Map A), and the higher the amount of task irrelevant information (i.e., Maps B and C), the more reasoning effort is needed when making decisions with these maps.

While our interviewed map-based decision making professionals did not use slope maps (Map D) for their daily work, they considered them as "nice-to-have", so these maps were also included in our study. The slope maps contain most task-relevant information in our tested maps. They show slope information explicitly on the map, and the respective information is explained in the accompanying legend. Thus, slope maps should be easiest to use for task domain novices. Slope was calculated in ESRI's ArcGIS and depicted in a diverging color scheme, employing the traffic light metaphor (green = go, red = stop). Slopes that are flat enough for a helicopter to land (i.e., below 14% steepness) are depicted with green shades, while slopes that are too steep for landing (i.e., above 14%) are shown in magenta shades (see Figure 30D). The amount of realism can be defined as a degree of verisimilitude with the real world (Zanola et al., 2009). I thus contend that shaded relief maps (Maps B & C in Figure 30) look more realistic than a contour map (Figure 30A), because contours cannot be seen in the real world.

Finally, the four tested map types also vary in graphic quality, or negatively put, in their degree of visual clutter. I quantitatively assessed this purely bottom-up vision concept in our test displays by means of the Subband Entropy clutter measure, proposed by Rosenholtz and colleagues (see section 2.3.3). Subband entropy was computed for the four map types shown in Figure 30, and I could find a robust order: There is most clutter in the slope maps (M=3.88, SD=0.06), followed by the dark hill shaded relief (M=3.51, SD=0.17), the light hill shaded relief (M=3.45, SD=0.16), and lastly, the contour maps (M=3.43, SD=0.18). The investigated factors are summarized in Table 3 below.

Table 3: Comparison of the map types used in Experiment III.

	contour map	shaded relief maps	slope map
degree of realism	(+)	(++)	(+)
depiction type (elevation information)	lines of equal elevation (absolute)	lines of equal elevation (absolute) & shaded relief (relative)	lines of equal elevation (absolute) & slope classes (absolute)
slope information type (amount)	implicit (+)	implicit (++)	explicit (+++)
visual clutter (Subband Entropy)	(+)	(++)	(+++)
reasoning effort	(+++)	(++)	(+)

The locations participants had to assess for potential helicopter landing were represented with black labels (numbers) on a yellow background to maximize saliency. Each stimulus contained six such locations for assessment. No other pieces of information (such as labels of place names) were contained in the map.

7.3 Procedure

Like the previous road selection experiment, this experiment took place in a lab equipped with standard personal computers connected to the Internet. The experiment was carried out digitally in a web browser displayed on a 17-inch computer screen set to 1280 x 768 pixel screen resolution. After filling out a background questionnaire, participants were then asked to assess slopes not steeper than 14%, on which a helicopter can land safely. To assure that all participants had the necessary background to complete the task, I first introduced them to the concept of slope and how it can be calculated. Participants then were shown how slope can be identified on a contour line map by using two pieces of information: the elevation information displayed with labels on the contour lines, and the ground distance information contained in the map scale bar. I then asked them to solve two warm-up tasks that were identical to the actual experiment and showed them the sequence of twelve maps described in the previous section (see the appendix of this thesis for a more detailed description).

As each participant was exposed to the same dataset, this experiment represents a within-subject experiment (Martin, 2008), in contrast to the between-subject design (shortest route / fastest route) in Experiment II. The reason that I chose this setup was that I did not vary the tasks in this Experiment III. Apart from that, the setup of Experiment III was similar to Experiment II. Like in the road selection experiment, participants were exposed to three time limits, which represent severe, moderate and generous time limits.

After a pilot testing procedure, in which 40 seconds was the average time pilot testers needed, the decision time limits were set to 20, 40 and 60 seconds. Thereby, I doubled limits compared to the road selection experiment. Like in Experiment II, time pressure was varied within all subjects. Time pressure was again simulated with a bar placed right of the map graphic, whose height was proportional to the seconds remaining and whose color was either red (for 20 seconds), orange (for 40 seconds) or yellow (for 60 seconds), representing different levels of time pressure. Subjects had to solve the slope detection task under all three time constraints and for all spatial display types described earlier. For each map stimulus, participants had to select the points which were at a slope which was flat enough by clicking a checkbox below the map. In other words, for each map stimulus six choices had to be made.

The order of the stimuli was systematically rotated to prevent learning biases due to potential ordering effects. For each map, participants had to select one or more locations that were flat enough for a helicopter to land, by clicking the respective checkbox below the map. For each map, six locations had to be assessed. The number of correct locations varied randomly from 1 to 5 per map. Overall, 50% of the labeled slopes were too steep to land a helicopter (and therefore wrong answers), and the other 50% were flat enough (and thereby correct answers).

After completing each task, participants were asked to rate their confidence of response on a scale from "1 – not confident at all" to "4 – very confident". Responses were collected digitally and included participants' accuracy (percentage of correct answers) as well as (self-reported) confidence as performance measures. The experiment took approximately 15 minutes to complete.

7.4 Collected Data: Analysis, performance measures and signal detection

The performance measures (accuracy and confidence) in this study are similar to the ones used in the previous road selection experiment (Experiment II). The confidence measurement is identical, as I used the same scale ranging from 1 (not confident at all) to 4 (very confident). However, it has to be mentioned that the response accuracy measurement differs from the road selection experiment: In the road selection experiment, I measured for each map stimulus whether the correct answer (shortest or fastest route) was chosen or not. In contrast, in the slope experiment, the selection had to be made for six different points for each map stimulus. Hence, the number of correct answers per map is a positive integer ranging from 0 to 6 and not just 0 or 1, like in Experiment II.

Finally, all selection choices were aggregated to obtain accuracy percentages per map stimulus. Accuracy of 100% means that all selection choices were correct, that is, every point at a slope steeper than 14% was correctly not selected, while every point at a slope flatter than 14% was correctly selected.

Using the conceptual framework of signal detection theory (SDT), response accuracy can be assessed with more analytical depth than when just comparing correct and false answers. SDT was originally developed for research on visual perception (Tanner and Swets, 1954), and has also been applied in geographic information science (Griffin and Bell, 2009). It can generally be employed for decision making under uncertainty, and especially when decisions have to be made based on two or more alternatives. In SDT, correct answers are coined "*hits*" or "*correct rejects*", and errors are called "*misses*" or "*false alarms*", respectively. This analysis framework can especially help identifying which kinds of errors participants might make, due to varying time constraints and spatial display types, and thus if errors might follow a particular pattern.

Applying this concept to this experiment, correctly selected locations per question are classified as *hits* (<14% steepness), and thus those (correctly) not selected locations are classified as *correct rejects* (>14%, see Table 4 below). Participant answers that are incorrect are classified as either *misses* or *false alarms*, respectively. A *miss* indicates a location that was not selected, even though it is correct (<14% steepness), and a *false alarm* occurs when participants incorrectly selected a location with a slope that is too steep (>14%). In other words, a miss is an overestimation of slope, while a false alarm represents the underestimation of a slope. Table 4 illustrates how these four possible types of responses can be classified within the data analysis framework of signal detection theory.

Table 4: Classification of correct and wrong answers according to signal detection theory.

	Reality	
Participants' decisions	Slope too steep (> 14%)	Slope flat enough (<14%)
Slope too steep (> 14%)	correct reject (true)	miss (false)
Slope flat enough (<14%)	false alarm (false)	hit (true)

7.5 Results

I first present the results regarding the context-related factor of time pressure, followed by the results for the map-related factor (map type), and the interaction of time pressure with map types on participants' accuracy and confidence ratings. Finally, I will briefly report on the effect of user-related factors on accuracy and confidence.

7.5.1 Effect of time pressure on response accuracy and confidence

Overall, participants' average accuracy values (see Figure 31) show a surprising, counterintuitive pattern: Participants are most accurate with the moderate time limit of 40 s (M=82.9%, SD=11.4%), followed by the most generous time limit of 60 s (M=72.2%, SD=15.2%), and lastly, as expected, the most severe time limit of 20 s (M=66.7%, SD=20.4%). Hence, there seems to be a speed-accuracy trade-off between 20 and 40 seconds, but no further trade-off between 40 and 60 seconds.

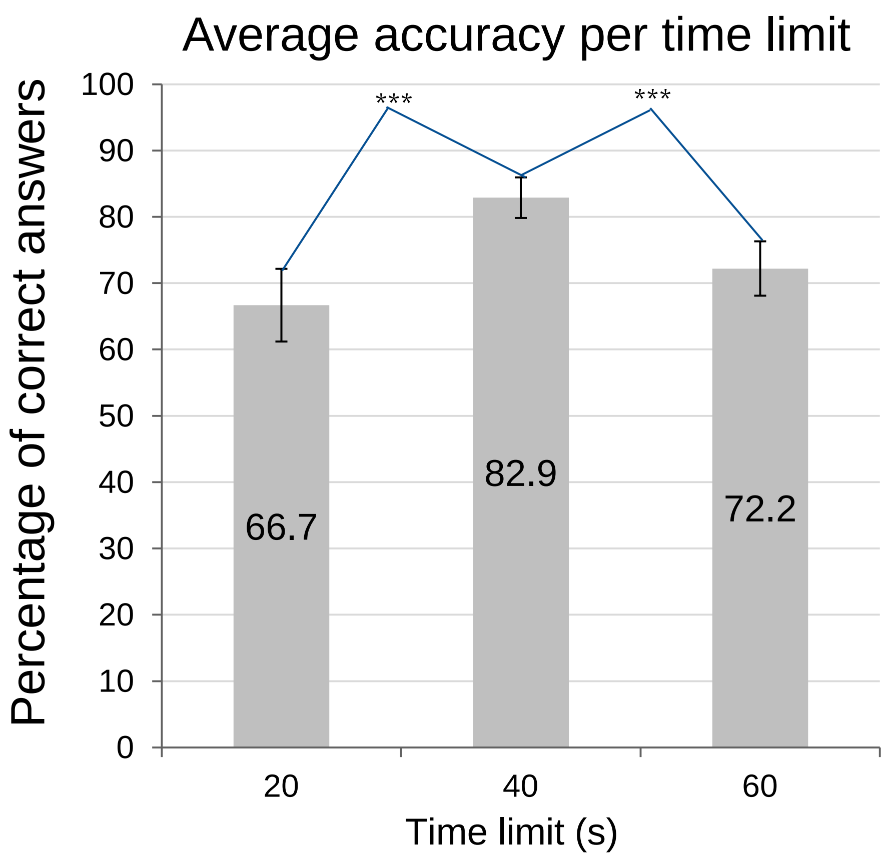


Figure 31: Average accuracy per time pressure limit. Error Bars: ± 2 SE, *** p< .001.

Similar to overall accuracy, participants' self-reported confidence (see Figure 32) is also highest for the moderate 40 s time limit (M=2.9, SD=0.4), followed by the most generous 60 s limit (M=2.7, SD=0.2) and, lowest, again as expected, for the most severe 20 s limit (M=2.7, SD=0.2). The performance increases from the 20 s time limit to the 40 s limit, and the performance decreases from the 40 s to the 60 s time limit are significant for both accuracy and confidence (p< .001). In analogy to the speed-accuracy trade-off, there also seems to be a speed-confidence trade-off between 20 and 40 seconds, but no speed-confidence trade-off between 40 and 60 seconds.

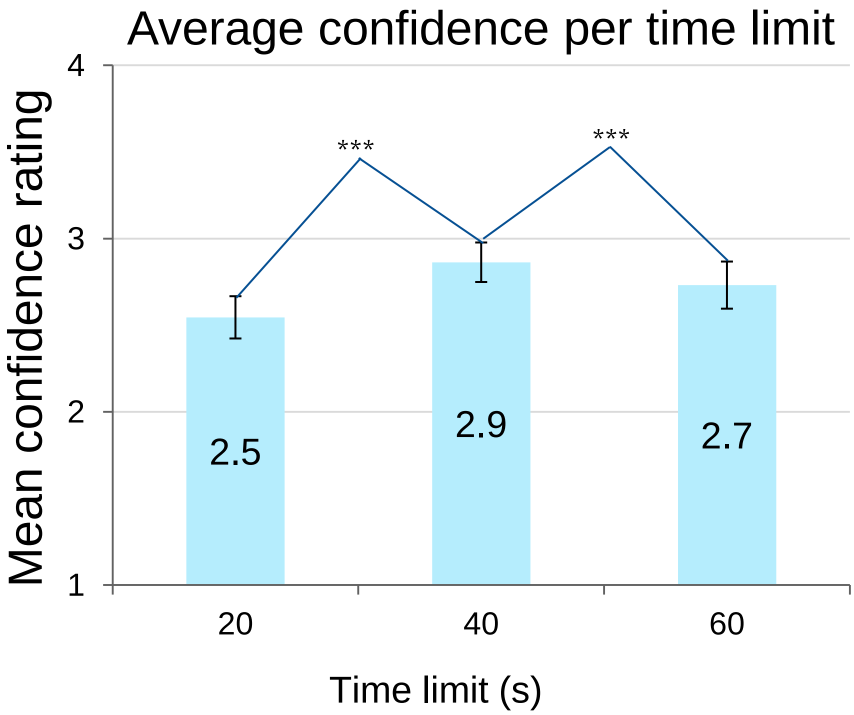


Figure 32: Average confidence per time pressure limit. Error Bars: ± 2 SE, *** p< .001.

7.5.2 Effect of map type on response accuracy and confidence

As expected, participants are significantly more accurate with the slope map (M=83.6%, SD=14.8%) compared to all other maps, as shown in Figure 33. However, accuracy is not better as predicted, but even worse with the shaded relief maps compared to all other maps. Participants' mean accuracy for the dark hill shaded relief map is 73.1% (SD=16.8%), and with 65.4% (SD=18.5%) it is lowest overall for the light hill shaded relief map. Somewhat surprisingly, participants perform even worse with the hill shaded relief maps that look more realistic, and contain more information than the most abstract contour map (M=73.5%, SD=13.0%). The difference between the contour map and the light hill shaded relief map is significant (p< .01), as well as the difference between the slope map and all other maps (p< .001).

As can be seen in Figure 34, in congruence with the accuracy response pattern, participants' confidence ratings are also highest for the slope map (M=3.1, SD=0.6), and higher for the contour maps (M=2.7, SD=0.5), compared to the lowest scoring hill shaded relief maps (dark: M=2.6, SD=0.5, light: M=2.5, SD=0.5). Surprisingly again, participants are significantly more confident in their performance with the contour map, compared to the light hill shaded relief map (p< .05) that contains more information than the contour map.

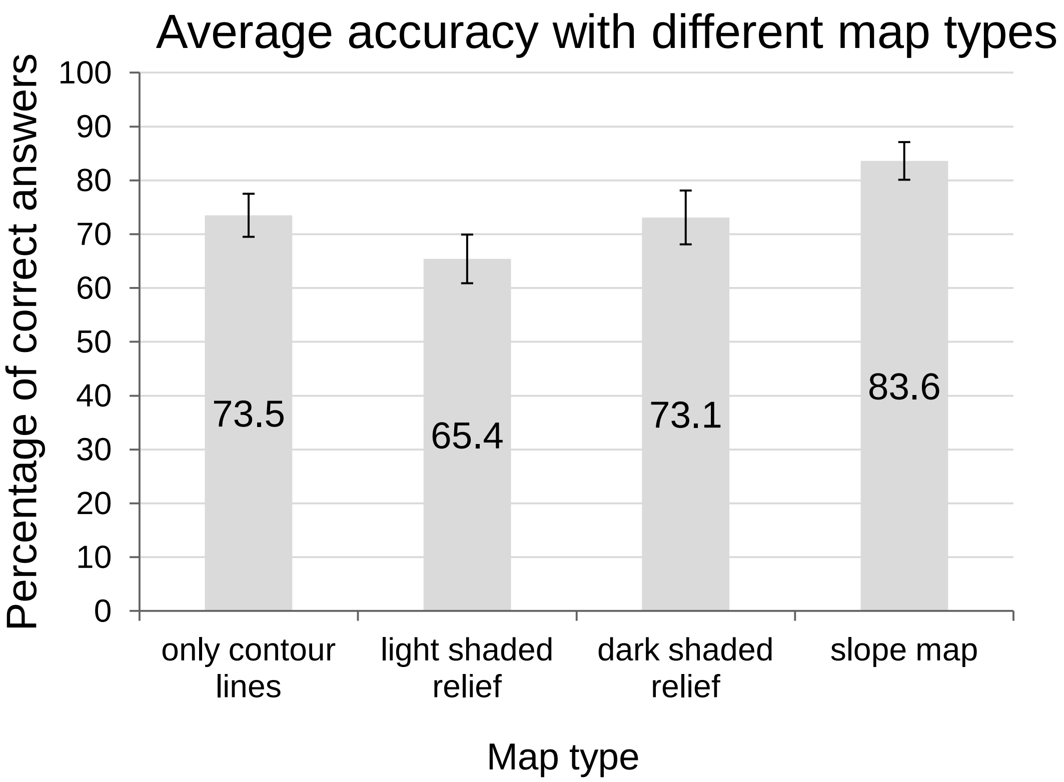


Figure 33: Average accuracy with different map display types. Error Bars: ± 2 SE.

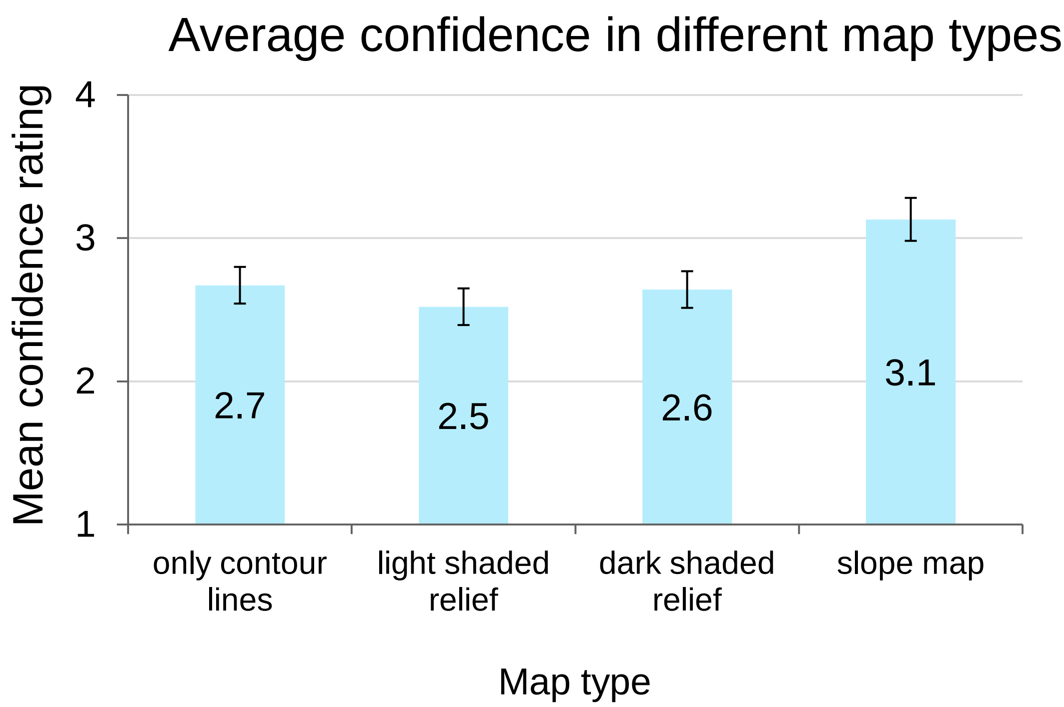


Figure 34: Average confidence in different map display types. Error Bars: ± 2 SE.

By using the framework of signal detection theory (see section 7.4), response accuracy can be analyzed in more detail. Overall, regardless of map type, misses (i.e., slope overestimation) occur more frequently than false alarms. Misses also occur more frequently than false alarms, independent of the tested time limits. For the correct answers, the "correct rejection" is overall more frequent than the "hit", also for all map types and all temporal conditions. As expected, the number of false alarms (i.e., slope underestimation) shown in Figure 35 is significantly higher for the light hill shaded relief map (M=2.0, SD=1.8) than for the dark hill shaded relief map (M=1.2, SD=1.2). In contrast, the number of misses (i.e., slope overestimation) is, again as expected, higher with the darker hill shaded relief map (M=2.7, SD=1.8) compared to the light shaded relief (M=2.1, SD=1.7). These differences are highly significant (false alarms: $p < .001$, misses: $p < .05$).

As shown in Figure 35, SDT provides additional insights on what kinds of decision errors might have specifically contributed to the unexpectedly low accuracy for the shaded relief maps. Similar to the other map types, participants seem to overestimate the steepness of the slopes more frequently with the dark hill shaded relief maps (i.e., higher number of misses) compared to the light shaded relief maps. Hence, a map with a lighter shaded relief might help reduce this potential source of error. However, one can also see in Figure 35 that one drawback of light hill shaded relief maps might be their relatively high rate of false alarms.

Figure 35: Average misses and false alarms with shaded relief maps. Error Bars: ± 2 SE. * p< .05, *** p< .001.

Signal detection theory provides additional measures (see Griffin and Bell (2009) for a detailed explanation and suggestions for applications to geographic information science), which can be calculated to further investigate participant accuracy. I will briefly discuss these measures in the remainder of this section.

To begin with, the hit rate (HR) measures the number of hits relative to the number of suitable slopes, while the false alarm rate (FAR) measures the number of false alarms relative to the number of slopes that were too steep. Both the HR (light 71.6%, dark 69.4%) and the FAR (light 23.3%, dark 14.2%) are higher for the light shaded relief maps. This seems plausible, as a high hit rate suggests that participants did *not* tend to overestimate the slope, while a high false alarm rate indicates that participants tended to underestimate the slope with the respective map.

Furthermore, other important SDT measures include sensitivity (d') and criterion (λ). The sensitivity or discriminability measure (d') is the standardized difference between the means of the "signal present" and "signal absent" distribution. In this experiment, the sensitivity measure is higher for the dark shaded relief maps (d' =1.6) than for the light shaded relief map (d'=1.3). This indicates that participants could distinguish between signals (suitable/flat slopes) and noise (unsuitable/steep slopes) more successfully with the dark shaded relief maps. The decision criterion (λ) measures the willingness of a respondent to say "signal present" in an ambiguous situation. A larger value of the criterion *"implies that the respondent requires stronger evidence before saying that the signal is present"*[13]. In the context of this experiment, "signal present" means that the slope is flat enough to land the helicopter. As the relief appears to be too steep with the dark shaded relief map, it is little surprising that the criterion is higher for the dark (λ=1.07) than for the light shaded relief map (λ=0.73).

7.5.3 Interaction of map type and time pressure

I now turn to the research question how certain map types might support participants in their decision making under varying time pressure scenarios.

Participants give most accurate answers with the (explicit) slope map under all time constraint conditions (see Figure 36). In the most severe time limit condition (20 s), participants score better with the most abstract contour map (M=71.2%, SD=30.5%), containing the least amount of information, compared to the more realistic looking shaded relief maps (dark: M=63.9%, SD=27.9% and light: M=53.3%, SD=40.1%). For this shortest time limit, the overall differences between contour map and shaded relief maps are significant ($p < .01$ for both shaded relief

[13] http://wise.cgu.edu/sdtmod/measures5.asp

maps). Accuracy scores generally increase from the most severe (20 s) to the moderate (40 s) time limit condition. The accuracy differences between contour maps and shaded relief maps are not significant in the moderate condition.

Overall, accuracy scores drop again for the highest scoring slope and contour maps under the least severe time constraint condition (60 s), while accuracy scores for the hill shaded relief maps do not change much for the 40 s and 60 s limit conditions. In other words, participants' accuracy with shaded relief maps only reaches the higher level of the other more abstract map types when participants are not under severe time pressure. Overall, the interaction effect between map type and time limits is significant ($p < .05$).

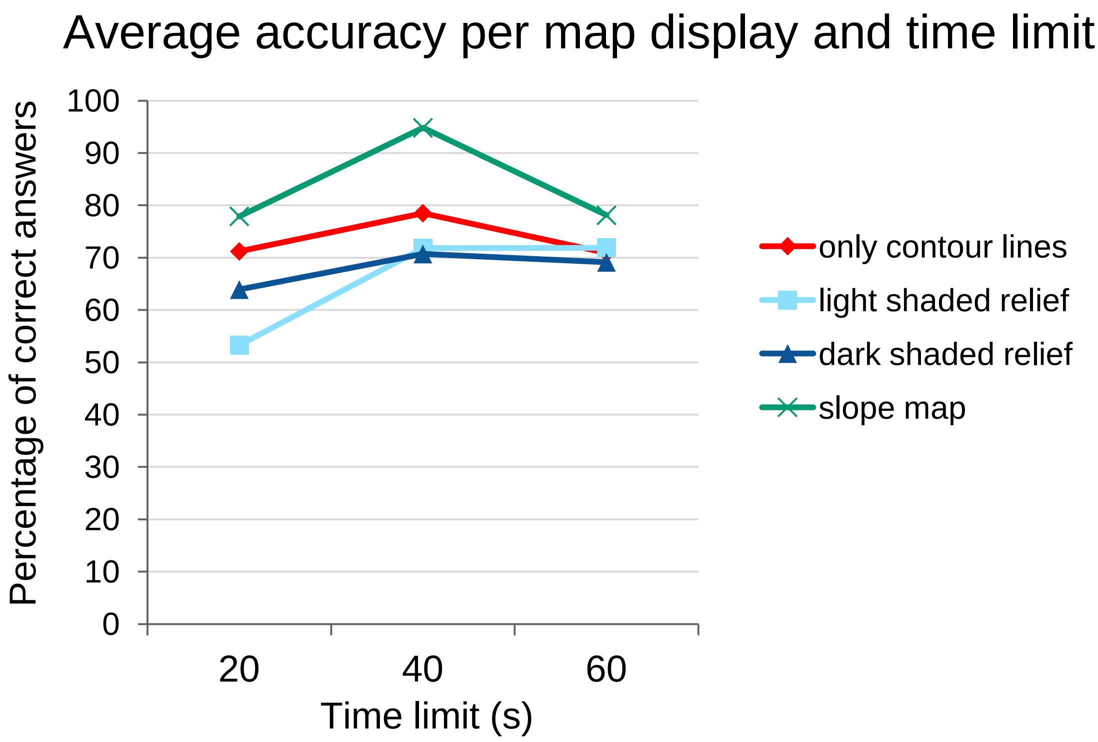


Figure 36: Average accuracy per map display and time limit.

A very similar response pattern can be observed in Figure 37, when looking at participants' confidence ratings. Again, mirroring accuracy scores, participants are most confident in their responses with the slope map, regardless of the given time limit. Participants' confidence is also consistently high with the contour line map. The difference between the average confidence ratings for the slope map and the shaded relief maps is only significant at the 20 s time limit. For this shortest time limit, the average confidence rating with the contour map is 2.6 (SD=0.1) and 2.3 (SD=0.1) with the shaded relief map. The rating difference between the contour map and both hill shaded relief maps is significant ($p < .001$). Only in the moderate time limit condition (40 s), confidence ratings for the hill shaded relief maps are higher than for the contour map. This is in contrast to participants' accuracy scores shown in Figure 36 earlier, as participants are more accurate with the contour map than with the shaded relief maps for the moderate time limit. Contrary to accuracy, the interaction effect between time limits and map type is not significant for confidence ($p > .05$).

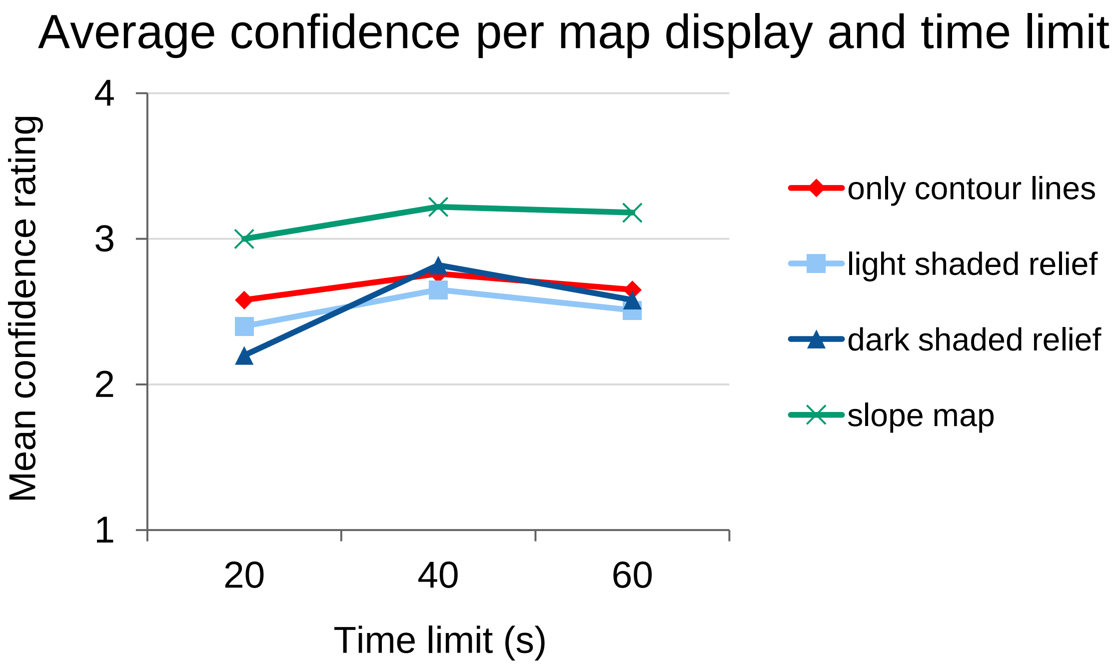


Figure 37: Average confidence per map display and time limit.

7.5.4 Effects of sex on response accuracy and confidence

Female response accuracy is again slightly – not significantly – better (M=75.8%, SD=12.7%) than male response accuracy (M=72.5%, SD=9.5%). Overall, male and female confidence do not differ significantly (M=2.5, SD=0.4 for both sexes). Under the short time limit of 20 seconds, males (M=2.6, SD=0.4) are again slightly more confident than females (M=2.5, SD=0.5), but the differences are not statistically significant (see Figure 38).

Finally, in this experiment accuracy and confidence are significantly ($p < .001$) and stronger correlated (Pearson's Rho = 0.45) among participants than in Experiment II.

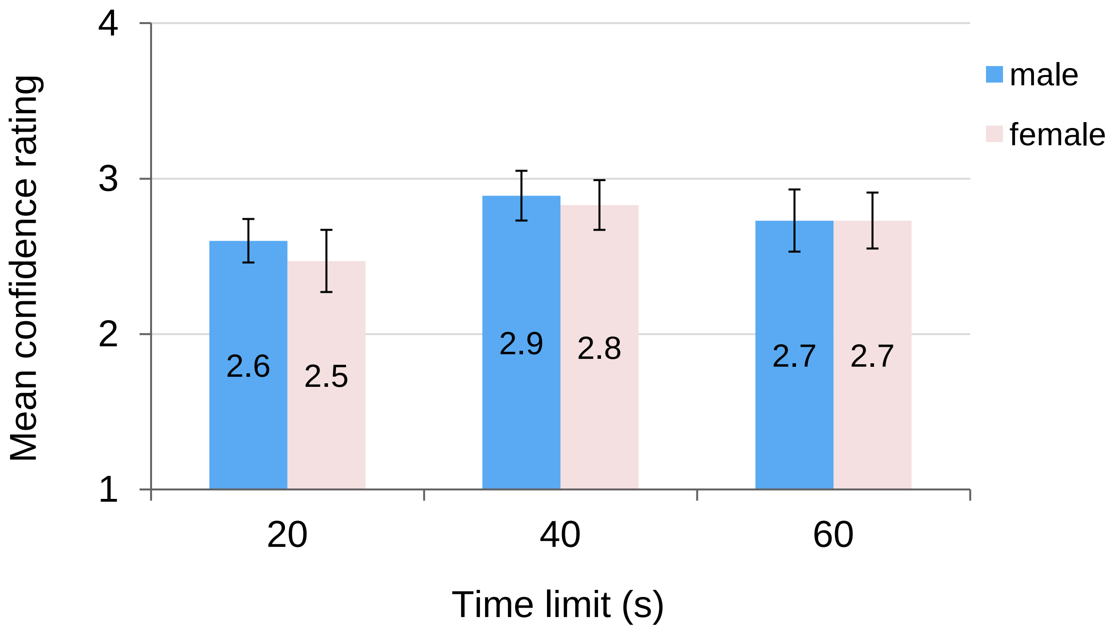


Figure 38: Confidence ratings grouped by sex and time limits. Error Bars: +- 2 SE.

7.6 Experiment III - Discussion

The main finding that both accuracy and confidence reach peaks when participants are under a moderate time limit seems especially counterintuitive. Both scores decrease significantly when participants have more decision time available. These results, which seem somewhat counterintuitive, do resemble the previously discovered inverted U-shaped response curve found by Hwang (1994) in a context which was not related to map-based decision making. One explanation could be that participants start to doubt in their decisions when they have more time. In Experiment II, there was neither an improvement in accuracy nor in confidence performance from a moderate to a generous time limit.

Based on Johnson et al.'s (1993) and Hwang's (1994) research results reviewed earlier, changes in speed-accuracy and speed-confidence trade-offs might be consequences of task difficulty. As both response accuracy and confidence decrease with a time limit more severe than 40 seconds, this slope detection task might have become significantly more difficult when participants had less than 40 seconds to respond. As a result, a clear speed-accuracy and speed-confidence trade-off effect can be found. Participant performance does not further increase from the moderate to the least severe time limit. Thus, the slope detection task is not getting easier with more available decision time beyond the 40 second time limit. In this experiment, the 40 second time limit seems to be the tipping point up to which time pressure actually increases performance. This pattern is in contrast to Experiment II, where no general time pressure effect on response accuracy can be found, even with overall shorter decision time limits, down to even 10 seconds decision time. One could argue that the road selection task on flat terrain is significantly less complex than a 3D slope detection task, and thus speed-accuracy and speed-confidence trade-offs are generally harder to find.

Regarding decision performance differences due to different map types, surprisingly, participants' accuracy and response confidence is unexpectedly low with the shaded relief maps. This result supports previous empirical work on hillshading maps (Philips et al., 1975; Potash et al., 1978) and studies by Hegarty and colleagues (Hegarty et al., 2008; Hegarty et al., 2009), who have shown that more realistic, 3D-looking displays while often preferred by "naïve users", do not necessarily increase performance. While three-dimensional shaded reliefs provide more task-relevant (but implicit) information, compared to the more abstract contour map, this increased information does not lead to more effective (accurate) or efficient (faster) decision making. One reason for this might, arguably, be that the implicit thematically relevant information is not presented in a cognitively and perceptually adequate and inspired way (Fabrikant et al., 2010; Swienty et al., 2008). While the shaded relief maps might contain more task-related information than the contour maps, they are also more cluttered (Rosenholtz et al., 2007), and thus might require more time for participants to visually parse. As Tufte (1983) would put it, the task-relevant *data-to-graphic-ink ratio* in the visuo-spatial display is not optimized for the task at hand. In fact, running a saliency model (Itti and Koch, 2001) on the stimuli, we find only one significant difference between the slope map and the other three tested map types (as shown in Figure 30). In the area along the bottom edge of the maps, where the density of the elevation contours is highest (i.e., the steepest area on the map), the slope map also shows darkest magenta shades between the contour lines. Moreover, the visual variable color hue does not seem to have much influence on this saliency map pattern, as running the saliency model on a gray scale version of the slope maps (i.e., removing color hue) yields an identical saliency map pattern. Another possibility why the more abstract maps could have performed better under time pressure is that our 3D maps with high graphic density might have a general relative disadvantage when shown at smaller screen sizes with lower spatial display resolution than the 2D maps.

However, participants are more accurate with the shaded relief maps compared to the more abstract contour line map when they have more decision time available, and also seem to be more confident in their responses with shaded relief maps when under less time pressure. In this case, participant accuracy and confidence seem to reflect participant preferences, when comparing results from this study with the results from Experiment I, in which more realistic

3D looking satellite image maps obtained higher preference ratings when participants have more decision time available. In other words, I did not find strong evidence for a "naïve realism" effect (Smallman and St. John, 2005), or over-confidence in more realistic maps in this experiment, as low accuracy scores co-occurred with equally low confidence ratings for the tested shaded relief maps. This could be due to the fact that the maps tested in Experiment III were not as realistic as the displays tested by Smallman and St. John. Another possible interpretation is that my participant sample consisted mainly of cartographic (design) professionals, and thus not "naïve" cartographers.

Not surprisingly, the 2D slope maps, containing most of the thematically relevant information, outperform all other map types with respect to effectiveness (i.e., accuracy) and efficiency (i.e., under all time limits), including participant confidence. In this case, in contrast to the shaded relief maps, the information increase has a positive effect on response accuracy and confidence, even though perceptually these maps appeared to be most cluttered (see Table 3). One reason for this could be that the thematically relevant information was communicated in a cognitively adequate (explicit), and perceptually salient way, using empirically validated cartographic design principles (Fabrikant et al., 2010). In other words, participants can be accurate and confident in their decisions even with an abstract (but computationally efficient) depiction method, but only when thematically relevant information is communicated explicitly and rendered in a perceptually salient manner. Although slope maps are not commonly known or used by map-based decision making experts under time pressure or the general public, participants and had no problem in detecting the relevant information without any training.

Finally, another explanation for the advantage of slope maps compared to the other map types might be the resolution of the contour lines. For instance, the interval of the contour lines in this experiment was 100 meters. Hence, the relevant slope information could not be obtained as precisely with the contour lines map as with the finer resolution of the slope map, which has a horizontal precision of 90m x 90m. On the one hand, one might argue that participants might have performed better with contour maps with a finer interval, such as 25 meters, because they would have more precise slope information. On the other hand, contour maps with a 25 meter interval would have been visually more cluttered and thus more difficult to parse. The relatively poor performance of shaded relief maps indicates that a more on information might not necessarily lead to better accuracy and confidence.

8. Experiment IV – Various tasks with an interactive map (virtual globe)

After having investigated how time pressure influences the accuracy and confidence in decisions made with static maps in Experiments II and III, this following Experiment IV sheds light on the research question of how time pressure influences the efficiency and effectiveness of spatial decisions when participants solve tasks using a fully-interactive virtual globe.

The aim of Experiment IV was to investigate to what extent human-map interaction is in accordance with interaction tool preferences, as assessed in Experiment I (see Chapter 4). I operationalized human-map interaction by measuring the usage of the four interaction tools that participants rated in Experiment I: zooming, panning, rotating and tilting. In Experiment I, participants stated that the 2D interaction tools, zooming and panning, are the most important interaction tools, regardless of time pressure. Furthermore, the preference to tilt a display was significantly lower when participants were under time pressure.

Another rationale for empirically investigating human-map interactions is the question of how interactivity might influence response accuracy and confidence in map-based decisions under time pressure. An open question in this context is to what extent interactivity actually enhances task performance of participants in a spatial task (i.e., accuracy and confidence). Studies discussed in section 2.4.1 have shown mixed results regarding the benefits of interactivity (James et al., 2002; Keehner et al., 2008; Marchak and Zulager, 1992).

This experiment consisted of four different tasks:

1. Elevation detection of two given points
2. Detection of the highest elevation on a given polyline
3. Selection of the steepest slope among three choices
4. Qualitative description of the terrain between two points

The motivation for using four different tasks was to explore how different complexity levels of typical map use tasks (see context-related factors in Chapter 1) influence human-map interactions, response accuracy and confidence. These four tasks were chosen because they all seem particularly relevant for decision making in the third dimension. Therefore, the usage of all interaction tools (zooming, panning, rotating and tilting) can lead to more accuracy. However, all tasks can also all be solved without interacting with the map, or, more precisely, without using any of the four interaction tools. Furthermore, the chosen tasks are similar to the ones participants solved in the three previous experiments, and are typical tasks in time pressure (e.g., emergency situation) and no-time-pressure conditions (e.g., excursion planning).

In the context of user-related factors, I specifically investigate whether people who are good at mental rotation actually rotate the display more, as previous work by Cohen and Hegarty (2007), and the preference ratings of Experiment I suggest. Other issues of interest include whether participants who rotate paper maps when navigating in the field (Lobben, 2004, 2007) will also rotate an interactive map more often, or how familiarity with video games might positively influence the efficiency and effectiveness of decision making with visuo-spatial displays, as shown by several authors (Feng et al., 2007; Terlecki et al., 2008). Moreover, I want to explore if experienced virtual globe users might also tilt the display more than virtual globe novices (Abend et al., 2011). Finally, I investigate if the patterns of higher male confidence in task performance found in Experiments II and III and several other studies (Furnham, 2001; Furnham et al., 1999; Lloyd et al., 2002) can be replicated in map-use tasks with an interactive map.

8.1 Participants

Twenty-one participants (11 male, 10 female) took part in this study. The majority were students and staff of the Geography department at the University of Zurich.

Two of the 21 participants mentioned they use maps “very frequently” (10%), five “frequently” (24%), eight “occasionally” (38%) and six “never” (29%) in their daily working lives. As for leisure time activities, three stated they use maps “very frequently” (14%), thirteen “occasionally” (62%), and five “never” (24%).

Participants were also asked how familiar they were with Google Earth or other virtual globes. Three participants mentioned they were “very familiar” (14%), eleven “rather familiar” (52%), five “rather not familiar” (24%) and two “not familiar at all” (10%) with virtual globes. It has to be stated that the persons who were “not at all” familiar with virtual globes were in fact both geographers by training. The distinction “non-geographer” and “geographer” can therefore not simply be made by measuring the variable “familiarity with virtual globes”.

Nine of our 21 participants (43%) mentioned playing video games at least occasionally, and fourteen participants (67%) mentioned they rotate paper maps when navigating in the field.

8.2 Materials

As this experiment focused on tasks where the third dimension is particularly relevant, I chose eight GPS tracks from mountain areas all over the world as stimuli. GPS tracks were downloaded from a GPS track sharing website[14] and were used as polyline overlays in Google Earth, the software which represents the most ubiquitous virtual globe (Schöning et al., 2008). Seven marker symbols were added to the GPS tracks. These marker symbols included one start location (using the standard red marker labeled "A" from the Google Earth icon collection), one end location (a red marker labeled "B" from the same icon collection), three points (red markers labeled 1, 2, and 3), and finally two pushpins (yellow and green). All other default layers in Google Earth (place names, points of interest, etc.) were deselected and thus not visible on the map. A screenshot of a sample stimulus with all the markers is shown in Figure 39.

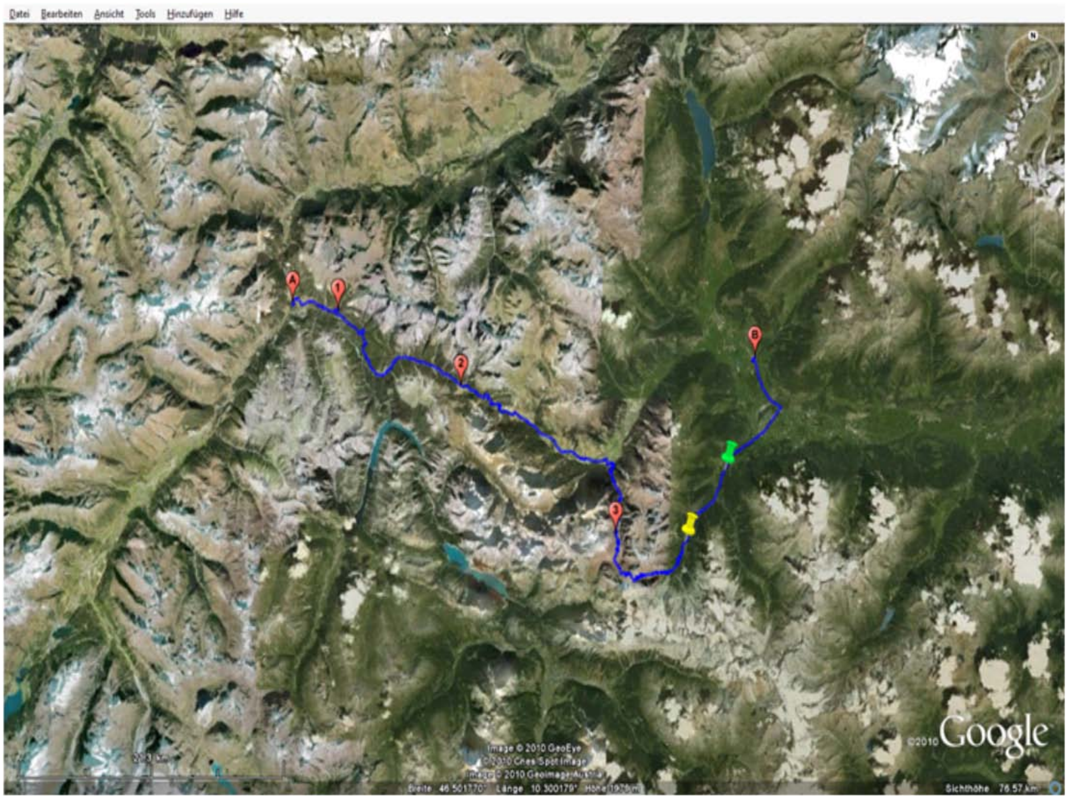


Figure 39: Screenshot of a GPS track including markers as used in Google Earth for Experiment IV.

8.3 Procedure

The experiment took place in a windowless office specifically designed for eye-tracking experiments. Participants were welcomed and asked to sign a consent form. Then, the experiment started with the Vandenberg Mental Rotation Test (MRT), as used in Experiment I, in order to assess participants' mental rotation abilities. The mental rotation test took approximately 15 minutes including warm-up.

[14] http://www.gpsies.com

Thereupon, participants were introduced to the computer-based part of the experiment. It was administered on a Dell Precision 390 Windows workstation, equipped with a 20-inch flat panel display set to a screen resolution of 1024 x 768 pixels. At the beginning of the experiment, participants were introduced to the four map interaction tools, zooming, panning, rotating and tilting, and other relevant map elements (such as the scale bar and elevation information) visible on the screen. The concept of slope and how to calculate it (rise over run) was explained to participants both by a printed sketch and in Google Earth. They were then asked to perform a warm-up task, which consisted of answering four questions on the basis of a sample GPS track. The four questions in the warm-up task were identical to the questions in the main experiment:

1. **Elevation AB:** What is the elevation (above mean sea level) of point A and point B?
2. **Highest point:** Where is the highest point (above mean sea level) along the entire path/GPS track?
3. **Steepest slope:** When travelling from A to B, at which of the three points (1, 2, or 3) does the path/GPS track have its steepest slope?
4. **Profile description:** How would you verbally describe the elevation profile between the yellow and the green pushpin? For instance, “only downhill”, “flat”, or “first uphill, then downhill”?

After the warm-up trials, the first half of the experiment started. In this first half, participants had to solve the four tasks four times, with four GPS tracks. After participants had solved all tasks for one GPS track, the geographical extent of the map changed, so that the full extent of the new GPS track could be seen on the screen. The sequence of the two scenarios (TP/time pressure, NTP/no time pressure) was systematically varied across participants: For one half of the experiment, participants were under a time limit of two minutes for solving all tasks for each track (TP = time pressure). Participants were instructed to answer all four questions within this temporal limit as accurately as possible. Like in previous experiments, this time limit was identified after pilot testing. For the other half of the experiment, participants were not given any time constraint at all (NTP = no time pressure), and they were told that they could take as much time as they needed for responding. The order of the eight GPS tracks was systematically varied as well.

At the end of the first half, participants were asked how confident they were in their decisions at each of the four different tasks. Confidence was assessed using a rating scale ranging from 1 (not confident at all) to 4 (very confident). Thereupon, the experiment continued with another set of four GPS tracks, and with the second condition (TP and NTP respectively). Finally, participants also had to rate their confidence about the second half.

After the computer-based main experiment, participants were finally asked to fill out a background questionnaire, in which they specified their map use experience, their familiarity with Google Earth and 3D representations, whether they rotated paper maps in the field, and how often they played video games. The overall duration of this experiment, including the Mental Rotation Test, was between 40 and 60 minutes, dependent on how quickly participants solved the tasks (especially for the NTP portion). After completing the digital portion of the experiment, participants were debriefed, and given a meal voucher for the university cafeteria in return for their participation.

8.4 Results

I first discuss how time pressure and the different tasks influence the frequency and the type of human-map interactions. Then, I report how response accuracy and confidence are affected by time pressure, with respect to the different tasks and interaction tools. Finally, I present how the user-related factors influence human-map interactions, response accuracy and confidence.

8.4.1 Effect of time pressure and task on human-map interactions

As mentioned in section 8.3, participants had to perform four tasks with four different GPS tracks under TP and NTP conditions. For every task, I recorded whether a participant used a tool or not. Accordingly, the maximum number of usages for each tool (zooming, panning, tilting and rotating) was four per task, and 16 (4 stimuli x 4 tasks per TP and NTP) overall per TP and NTP conditions.

The average frequencies of human-map interactions per task and interaction type are shown in Table 5. When comparing the tables, three findings stand out particularly: Firstly, as hypothesized, each tool was used more frequently when participants are not under time pressure (see "overall" row sums in Table 5). Secondly, in each of the four tasks, participants on average interacted more with the map when not under time pressure. In Table 5, the cell values represent average interaction tool frequency, with standard deviations in brackets. The maximum possible value for each cell is 4, and 16 for the overall row and column sums as mentioned earlier. The signs after the brackets indicate that the tool is used either significantly more (++, $p < .05$), more (+), less (-), or significantly less (--) than the tool average per task/column. The table also reveals that, overall, participants interacted most for the *steepest slope* task, the most complex of the closed-ended questions, and least for the *elevation detection* task, what can be considered the easiest task in the experiment.

As hypothesized, panning and zooming are the most frequently used tools, followed by tilting and rotating. This order is consistent for both TP and NTP conditions. On average, participants pan more than they use all other interaction tools, for both conditions and all tasks. There is only one example (*highest point,* TP) when people use one of the 3D interaction tools (i.e., tilting) more than a 2D tool (in this case, zooming). However, the sequence is dependent on task complexity. Only for the highest point identification task, people used one of the 3D interaction tools (i.e., tilting) more than a 2D tool (i.e., zooming).

Participants do not only tilt more frequently, but also zoom and pan more without time pressure ($p < .05$ for TP/NTP differences for tilting, zooming and panning). For rotating, the differences between TP/NTP are not significant ($p = .057$).

Table 5: Average frequencies of human-map interactions under time pressure (TP) and without time pressure (NTP). Standard deviations in brackets.

Time pressure (TP)

	Elevation AB	Highest point	Steepest slope	Profile description	*Overall*
Zooming	1.6 (1.7) +	1.6 (1.3) -	2.5 (1.5) -	2.1 (1.5) -	***7.8 (4.6) +***
Panning	1.7 (1.7) +	2.6 (1.4)++	3.6 (1.0)++	3.4 (0.9)++	***11.4 (3.7)++***
Rotating	0.3 (0.7)--	1.0 (1.5)--	2.2 (1.5) -	1.7 (1.3) -	***5.2 (3.7)--***
Tilting	0.4 (0.9)--	2.0 (1.8) +	2.4 (1.6) -	1.7 (1.4) -	***6.4 (3.7) -***
Overall	***4.0 (4.1)***	***7.2 (5.0)***	***10.7 (3.8)***	***9.0 (2.9)***	

No time pressure (NTP)

	Elevation AB	Highest point	Steepest slope	Profile description	*Overall*
Zooming	2.8 (1.8)++	2.8 (1.4) +	3.2 (1.2) +	2.4 (1.4) +	***11.2 (4.6) +***
Panning	2.8 (1.7)++	3.2 (1.2)++	3.9 (0.3)++	3.5 (0.8)++	***13.4 (3.3)++***
Rotating	0.2 (0.9)--	2.1 (1.3) -	2.5 (1.2)--	2.1 (1.4) -	***7.0 (3.4)--***
Tilting	0.2 (0.9)--	2.4 (1.5) -	3.1 (0.9) -	2.0 (1.2)--	***7.7 (2.9)--***
Overall	***6.0 (4.1)***	***10.6 (4.3)***	***12.7 (2.1)***	***10.0 (2.4)***	

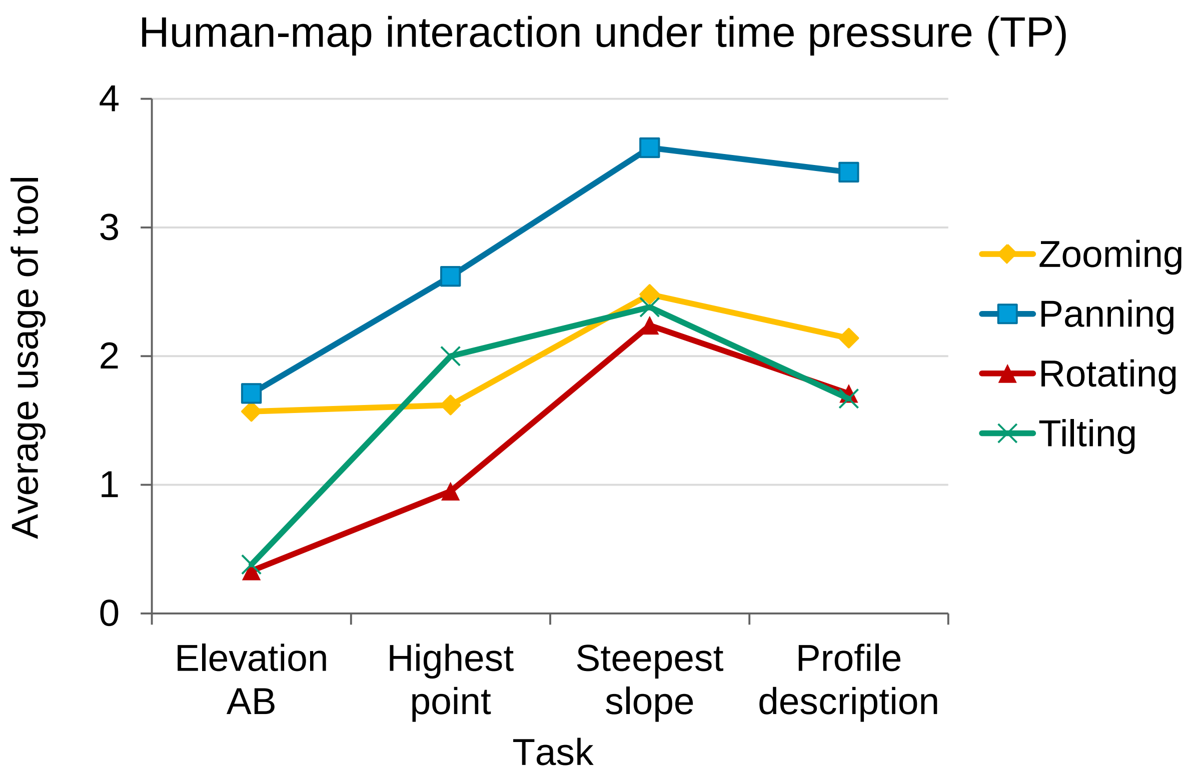


Figure 40: Average usage of interaction tools under time pressure (TP).

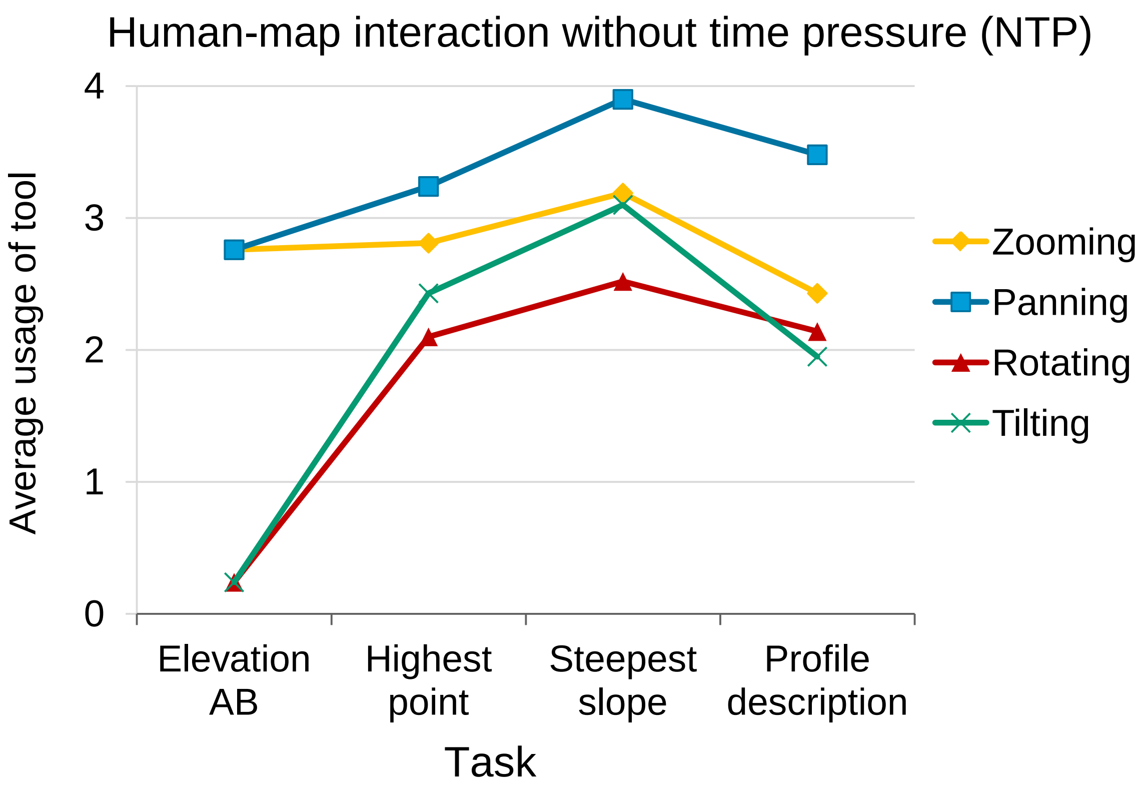


Figure 41: Average usage of interaction tools without time pressure (NTP).

The interactions between decision time, task and the frequency of human-map interactions (i.e., tool usage) are summarized in Table 6. In this table, the shaded cells containing "--" indicates that the interaction tool is used significantly less ($p< .05$) under time pressure (TP), than without time pressure (NTP), "-" that the tool is used less under time pressure (but not significantly so), and "+" that the tool is used more under time pressure. The results suggest that it is both task- and tool-dependent how an increase in decision time actually influences the intensity of human-map interactions. The differences in interaction intensity between TP/NTP conditions are most striking for zooming: Participants zoom significantly more under NTP in all tasks except *profile description*. As regards the other tools, participants pan significantly more for *elevation AB* and *highest point*, rotate significantly more under time pressure only in the *highest point* task, and finally tilt significantly more in the *steepest slope* task.

For each tool, there is at least one task where the tool is used significantly more when people have more decision time. Comparing the tasks, the TP/NTP differences are most prominent for the two tasks *highest point* and the *steepest slope*.

Table 6: The tensity decrease of human-map interactions per task and tool under time pressure.

	Elevation AB	Highest point	Steepest slope	Profile description
Zooming	--	--	--	-
Panning	--	-	-	-
Rotating	+	--	-	-
Tilting	+	-	--	-

8.4.2 Effect of time pressure and task on response accuracy and confidence

Next, I investigate how time pressure and the differences in task complexity influence response accuracy and confidence. The analysis of accuracy focuses on the two tasks *highest point* and *steepest slope*. The rationale for this is that response accuracy in the other two tasks is either a question of how detailed participants zoom in (elevation AB), or it cannot be measured precisely (profile description). Therefore, analyses of accuracy seem little meaningful for those tasks.

For the *highest point* task, I computed the deviations (i.e., absolute values) of participants' answers from the true highest point for each of the eight tracks, and aggregated them for both

TP and NTP conditions. For the *steepest slope* task, I recorded whether participants selected the correct slopes or not, and then grouped responses for each of the conditions.

For the *highest point* task, the average deviation from the true elevation is 421.3 meters (SD=293.1) in the TP, and 284.3 meters (SD=320.7) in the NTP condition. As expected, average response accuracy is worse under time pressure for this task. However, the TP/NTP differences are statistically not significant. Selecting the steepest slope, participants on average answered 1.52 (out of 4 possible) answers (SD=0.75) correctly in the NTP scenario, while under TP the number of correct answers is slightly lower (M=1.48, SD=0.93). The TP/NTP differences are again not significant for this task.

Confidence was measured on a scale from 1 (lowest) to 4 (highest) for each task. As Table 7 shows, participants on average are most confident solving the *elevation AB* task in both time limit conditions, and least confident at the *steepest slope* task. Average confidence values are higher for NTP for all tasks, as expected. The average of all confidence ratings for the NTP condition (M=3.0, SD=0.4) is significantly higher than for the TP condition (M=2.8, SD=0.5) with $p < .05$. Regarding each of the four tasks separately, however, the differences are significant only for the *steepest slope* task. At this task, the confidence ratings are generally lowest among all tasks under time pressure.

**Table 7: Mean confidence ratings per task and TP/NTP conditions.
Standard deviations in brackets. * p < .05**

	Elevation AB	Highest point	Steepest slope	Profile description	***Overall***
TP	3.7 (0.5)	2.4 (0.8)	1.8 (0.9)	3.2 (0.8)	***2.8 (0.5)***
NTP	3.8 (0.4)	2.6 (0.7)	2.7 (0.7)*	3.4 (0.6)	***3.0 (0.4)****
Overall	***3.7 (0.4)***	***2.6 (0.6)***	***2.1 (0.8)***	***3.3 (0.6)***	

In summary, time pressure seems to have a stronger effect on response confidence than on accuracy when people make map-based decisions with a virtual globe. People's confidence seems to be most negatively affected by time pressure in the most complex task of detecting the steepest slope.

8.4.3 Effect of interactivity on response accuracy and confidence

To investigate the potential interaction between interactivity and accuracy, I conducted a Pearson correlation analysis between the usage frequencies of each interaction tool and the accuracy measures for the two tasks *highest point* and *steepest slope*. This analysis shows only one significant correlation (out of 16 possible): For the *steepest slope* task, the quantity of zooming is positively correlated with response accuracy (Pearson's Rho = 0.46, $p < .05$) in the

NTP condition. This might suggest that zooming leads to a significantly higher accuracy for the slope detection task, particularly when participants are not under time pressure. However, it could also imply that people who are better at interpreting slopes tend to zoom more. Under time pressure, the frequency of zooming and response accuracy are also positively correlated (Pearson's Rho = 0.33), albeit not significantly ($p > .05$).

I also calculated the correlation between human-map interactions and confidence. Human-map interactions significantly influence the confidence with which people make decisions in three (out of 32) possible cases. All these cases concern the task *elevation AB*: Under TP, there is a negative correlation between rotating the display and confidence (Pearson's Rho = 0.52, $p < .05$), while for NTP both rotating and tilting (Pearson's Rho = -0.57, $p < .01$ for both tools) are negatively correlated with confidence. This could imply that people who are less confident might rotate and tilt the display more, and does not necessarily suggest that rotating and tilting would lead to a higher confidence in responses.

8.4.4 User-related factors

The only statistically significant effect of user-related factors on the performance measures is that people who play video games on a regular basis are also more confident in their decisions with interactive globe viewers. On average, "video gamers' " (N=9) confidence ratings are 3.2 both for TP and for NTP (SD=0.4 in both cases), compared to "non-video-gamers' " (N=12) average values of 2.5 for TP (SD=0.4) and 2.9 for NTP (SD=0.3). A one-way ANOVA confirms that this effect of familiarity with video games on response confidence is significant for both TP ($p < .01$) and NTP ($p < .05$). Video gamers are more confident than non-video-gamers in all tasks. However, these differences are only significant for the *steepest slope* task ($p < .05$). "Video gamers" are also more accurate than "non-video gamers" in the *steepest slope* task for both temporal conditions (overall correct answers M=3.2, SD=1.5, M=2.9, SD=0.8) and in the *highest point task* (video-gamers' deviation from the correct value: M=580.9, SD=379.1, non-video gamers' deviation: M=799.2, SD=465.9), albeit in both cases not significantly.

The results do not suggest, as previously hypothesized, that people who rotate a paper map when navigating in the field also rotate an interactive display more than "non-rotators": On average, "non-rotators" (N=7) even rotated the interactive display more than the "rotators" (N=14), in both the TP (non-rotators: M=8.4, SD=4.7, rotators: M=5.4, SD=2.9) and the NTP (non-rotators: M=8.4, SD=2.4, rotators: M=7.4, SD=3.1) condition.

The average Mental Rotation Score of participants was 20.7 (SD=8.0). Females (N=10) on average scored 15.7, while males (N=11) scored 25.2 on average. The median was 21.0. This performance is not significantly different from participants of Experiment I. Overall, the results do not clearly support that generally high-spatial or low-spatial participants rotate the map more. On the one hand, low-spatial participants (N=10) rotate the interactive display more under time pressure. On the other hand, high-spatial participants (N=11) rotate more frequently without time pressure. The overall differences are, however, statistically not significant.

Finally, as hypothesized, there is no clear evidence that participants who are more familiar with virtual globes would tilt the display more often: On average, participants who were less familiar with virtual globes (N=7) tilted more than experienced virtual globe users (N=14) in both temporal conditions.

I again find that male (M=11) confidence is higher than female (M=10) confidence, like in Experiments II and III. Male confidence is higher under time pressure in each of the four tasks, but in only two out of four tasks without time pressure. In both tasks where response accuracy was measured, males are on average also more accurate than females under time pressure.

8.5 Experiment IV - Discussion

Usage frequency rankings of the four interaction tools in this experiment are in accordance with the patterns of the preference ratings in Experiment I: Zooming and panning are the most frequently used interaction tools under and without time pressure, and used significantly more than rotating and tilting, even though I specifically assessed tasks where the third dimension is relevant. Overall, these experimental results on interactivity confirm the view of Harrower and Sheesley (2005) that the 2D interaction tools zooming and panning are the key component in any information visualization and are more important than the 3D tools rotating and tilting.

Another way to interpret these results is regarding interacting with a display as cognitive costs (Bleisch, 2011; Nielsen, 2007; Shepherd, 2008), which have to be invested in order to realize a certain benefit. A user has to decide for which task and context cognitive demands have to be invested in order to realize an inference-making or decision-making benefit (Smith et al., 1982). In the context of this experiment, the benefit can be measured in a high response accuracy and confidence, and under time pressure people can only spend a limited amount of cognitive costs. From this cost-benefit point of view, it seems that users are more willing to invest the cognitive costs to zoom and pan, because they believe that the benefits of zooming and panning justify the effort made while performing these interactions.

Furthermore, the analysis of human-map interactions demonstrates that people generally interact less with a virtual globe when they are under time constraints. This is statistically significant for all four interaction tools. Thus, time pressure indeed seems to have an effect on human-map interactions. The TP/NTP differences were most striking with respect to zooming: Participants zoomed significantly more often without time pressure in three of four possible tasks. One interpretation is that people regard the benefit-cost ratio of zooming as higher when they are not under time pressure.

In the map use preference experiment (Experiment I), participants stated they would tilt the map more without any time pressure for a road selection task. However, in the experiment reported in this chapter, in only one out of four tasks participants tilt actually significantly more in the NTP condition (*steepest slope).* In this task, *steepest slope*, firstly, overall confidence ratings are lower than in all other tasks, and, secondly, confidence significantly decreases under time pressure. Regarding confidence as a performance measure (see Chapter 3) and taking into account the work of Hwang (1994) – who made the case that performance decreases under time pressure might be a manifestation of a high task complexity – these findings suggest that *steepest slope* might be the most complex of the four tasks. This, in turn, could imply that the benefit-cost ratio of tilting is best for solving complex map-based tasks without time pressure.

When detecting the elevation of two points (a task where response confidence is generally high), participants hardly rotate or tilt at all. A lower time pressure does not lead to higher levels of rotating or tilting the display for this task either. One interpretation of this could be that identifying the elevation of two points represents a relatively simple task, for which the 2D interactions panning and zooming are sufficient to get an accurate answer. In contrast, the benefits of using the 3D interaction tools, rotating and tilting, do not seem to outweigh the cognitive costs for these interactions.

There is no strong evidence for any speed-accuracy trade-off in this experiment. For the complex task of detecting the steepest slope, response accuracy is even higher under time pressure. This might be another indication for the positive effect of time pressure on decision making shown by other authors for non-map-related tasks (Hwang, 1994), which has been found also for map-based decision making in my slope detection experiment (Experiment III). Interestingly, however, the *steepest slope* task is the only one among the four tasks where users actually are significantly more confident when not under time pressure. This is another example for the discrepancy between response accuracy and confidence, which was also evident in Experiment II.

Overall confidence results replicate the speed-confidence trade-offs found in Experiments II and III. This indicates that spatial tasks do not differ from non-spatial tasks (Maule, 1998; Maule and Andrade, 1997; Smith et al., 1982), as far as the speed-confidence trade-off is concerned. Thus, the speed-confidence trade-off might not only be a useful characterization of human map-based decision making with static maps, but also hold true for interactive maps.

There is one particular task where interacting with the map actually increases the *effectiveness* of map-based decisions: Participants who zoom in and out of the display are more accurate in the *steepest slope* task. This positive effect of interactivity is statistically significant when participants are not under time pressure, and supports the findings of studies in the fields of visual object recognition (James et al., 2002) or acquiring spatial knowledge of a virtual environment (Peruch et al., 1995), suggesting that interacting with spatial displays can actually help making people more accurate decisions. In other words, the results suggest that the benefits for zooming might be high enough to actually motivate the effort made to interact.

On the contrary, this does not seem to be the case for the 3D interaction tools, tilting and rotating: There is no evidence that the benefits of the 3D interaction tools are enough to justify the costs of interacting. These advantages – in terms of a higher benefit-cost ratio – of the 2D tools over the 3D tools are in accordance with the actual frequencies of human-map interactions (regardless of accuracy), as participants also perform the 2D interactions – such as zooming – more frequently than the 3D interactions rotating or tilting. While Fabrikant and Lobben (2009) have argued that "*what people think they want is not always what is best for them*" (p. 141), the interactive map users in this experiment actually *do* seem to perform the interactions that *are* best for them.

The only significant effect of user-related factors is that "video gamers" are more confident in their map-based decisions. Differences in confidence between video- and non-video-gamers are especially significant under time pressure. This might indicate that playing video games especially enhances *efficient* map-based decision making. Thus, playing video games might not only have a positive effect on spatial abilities (Feng et al., 2007; Terlecki et al., 2008), but also on efficient map-based decision making.

While the results also suggest that video gamers' response accuracy in spatial tasks is higher, this effect is statistically not significant. This pattern of "significant differences in confidence, but no significant differences in accuracy" between "video gamers" and "non-video gamers" resembles the differences between males and females found in Experiments II and III. As seven out of our nine (78%) "video gamers" are males and eight of the twelve (67%) "non-video-gamers" are female, there are certain interactions between sex and video game familiarity that have to be born in mind. On that account, the correlation between video game experience and response confidence might be a spurious one.

However, sex and other individual and group differences, such as mental rotation abilities and familiarity with virtual globes, influence neither people's confidence nor response accuracy in interactive map-based decisions significantly: People who claim to rotate maps in daily life overall do not use rotation tools in virtual globes more than "non-rotators", and experienced users of Google Earth do not tilt the display more than virtual globe novices. The results do not suggest that "high-spatial" participants (i.e., those who score high on the mental rotation test) overall rotate the display more often than "low-spatial" participants. This might also imply that the preference for the rotating tool, as shown in Experiment I, is more affected by spatial abilities than the actual human-map interaction via the rotating tool is. As the rotating tool does not, in most cases, influence the visibility of the task-relevant information, this result is in concert with findings of Keehner (2007), who suggested that interacting with the display is actually less important than seeing the task-relevant information.

9. General discussion

In the previous chapters, I have presented the results of my empirical work on map-based decision making under time pressure. In this chapter, I will discuss the main findings and their relevance for the broader context of this thesis. Besides the context-related factor of time pressure, I will also consider how map-related factors (spatial display types, map design and interactivity) and user-related factors (i.e., users with different backgrounds) influence the effectiveness and efficiency of map-based decision making. Therefore, this discussion is divided into the different influencing variables *decision time, spatial display*, *interactivity*, and *user characteristics,* as introduced in Chapter 1. The final section of this thesis is devoted to an overview of strengths, weaknesses, and implications of the four experiments.

9.1 Decision time (Time pressure and time limits)

Previous studies (Dillemuth, 2009; Kuoa et al., 2006) have shown that response time (as a dependent variable) for decision making tasks is affected by different map-related factors. In my experiments, I found that decision time limits (as an independent variable) also seem to affect map-based decision making in a number of ways.

My experimental findings indicate that there are several spatial display types and interaction tools whose suitability is very much dependent on time pressure. In contrast, other maps and tools seem to be equally suitable both under and without time pressure (see also sections 9.2 and 9.3). Overall, these findings suggest that certain maps might support *effective* (i.e.., accurate), but not *efficient* (i.e., accurate **and** fast) decision making (Coors et al., 2005; Dillemuth, 2005; Smallman et al., 2001).

The findings on map use performance under time pressure support the view that short response time limits lead to suboptimal (that is, not perfectly accurate) decisions. Considering previous work in decision making about bounded rationality (Simon, 1959), one possible interpretation of this is that time pressure can indeed be regarded as another factor that impairs optimal map-based decision making.

When selecting the shortest or fastest routes, participants are only slightly (and not significantly) less accurate when having less than 10 seconds time to respond. This implies that a speed-accuracy trade-off might exist in road selection tasks, but this trade-off does not seem to be as distinct as in other fields outside cartography (Maule and Edland, 1997; Wickelgren, 1977). Based on studies by Johnson et al. (1993) and Hwang (1994), those changes in speed-accuracy trade-offs might be consequences of task difficulty. Following these assumptions, the relatively weak speed-accuracy trade-off could be due to the fact that road selection is not a particularly complex task compared to other decision making contexts outside cartography.

In contrast, participants are significantly more confident in their road selection choices when having more than 10 seconds response time. Therefore, the speed-confidence trade-off, found in research outside cartography (Maule, 1998; Maule and Andrade, 1997; Smith et al., 1982), seems to be more striking than the speed-accuracy trade-off for road selection tasks. This in turn indicates that the effect of time limits on response accuracy might be less important than people think. Furthermore, both the speed-confidence and the speed-accuracy trade-off are more significant between a short and a moderate time limit than between a moderate and a generous time limit. This suggests that the road selection task is getting more difficult under time limits shorter than 20 seconds, but that the task is not substantially easier when having more than 20 seconds response time.

While time pressure does not significantly affect response accuracy – even with shorter time limits – in the road selection experiment, there is a clear speed-accuracy trade-off between the severe (20 second) and moderate (40 second) time limit for the slope detection task. This might indicate that slope detection is indeed a more difficult task than road selection, and thus requires more (than 20 seconds) response time to be solved with high accuracy and confidence.

Somewhat surprisingly, participants are on average significantly less accurate and confident with more available decision time in this slope detection task, more precisely, when they have more than 40 seconds response time. In this study, the tipping point to which time pressure actually increases performance seems to be in the vicinity of 40 seconds decision time, while a more generous time limit results in less overall accuracy and confidence. This pattern of the "inverted U-shaped curve" might seem counterintuitive, but it has also been found in decision making studies not related to maps (Hwang, 1994). It might be a result of certain positive effects of time pressure, such as a greater task involvement and the need to work harder when a deadline is imposed (Maule et al., 2000), or an optimal amount of arousal that leads to the best performance (Yerkes and Dodson, 1908). This pattern might also be due to the specific time limits used in the experiment, as most of the contents of working memory are lost after about 30 seconds.

Time pressure does not only influence response accuracy and confidence, but also how people interact with visuo-spatial displays. When solving tasks with a virtual globe, three of the four interaction tools are used significantly more when participants are not under time pressure.

Furthermore, time pressure does not significantly affect response accuracy, but it affects confidence in interactive map-based decisions. This replicates the pattern found in the road selection experiment with static maps, where the speed-confidence trade-off was more striking than the speed-accuracy trade-off. These results provide more evidence for the phenomenon that time limits seem to have a greater effect on response confidence than accuracy, both with static and with interactive maps.

In summary, the effect of time pressure on the efficiency and effectiveness of map-based decision making seems to be very task-dependent and unclear. However, the experimental results allow for the following main finding on the effect of time pressure: ***When making map-based decisions, time pressure seems to affect human confidence more than accuracy.*** This resembles findings from Pew (1969), who has found logarithmic increases in confidence co-occurring with only linear increases in response time, in a study not related to maps, but on detecting light signals. Before generalizing these results, one has to bear in mind that the response time limits used are probably critical for response accuracy: With a time limit that is not sufficient to solve a task properly, response accuracy will probably suffer significantly in every experimental context.

9.2 Spatial display factors (Display types and design)

Not surprisingly, road maps seem to be the most preferred display type for road selection tasks under time pressure, as the task-relevant information (the road network) is represented in a cognitively adequate and perceptually salient way (Fabrikant and Goldsberry, 2005; Fabrikant et al., 2010; Swienty et al., 2008). Therefore, road maps might require less reasoning effort. In contrast, the aesthetic details (such as the relief depiction or a panoramic view) in the more realistic displays are regarded as irrelevant when selecting roads under time pressure. It can be assumed that map design is highly relevant for map use preferences, as previous work (Gill, 1993) has shown that the design of different road maps might also influence the preference ratings.

For time-critical decision making contexts such as road selection, satellite images do not seem to be a suitable spatial display type, most probably because they do not contain the relevant road network information. People seem to regard these maps as more suitable for situations where there are no short decision time limits. One possible explanation for this – which was also stated verbally by participants – is that satellite images provide a good overview and depict the area in more detail than road maps do. These aspects seem less relevant when making map-based decisions under time pressure, when the relevant information has to be retrieved quickly from the map. Not only participants in controlled experiments, but also experts in the field of map-based decision making seem to prefer familiar displays when under time pressure. Altogether, map users seem to be more open towards using novel spatial displays for decision making without time pressure. This seems plausible, as under severe time pressure there is no time for decision-makers to familiarize with a novel representation or interaction method. What matters more in this context is quickly extracting the most task-relevant information.

Being under time pressure when making map-based decisions might lead to filtration or omission of irrelevant information on the map (Miller, 1960). If irrelevant details are not displayed on the map – as is the case in more abstract and less realistic representations – less cognitive effort is needed to parse the map. This abstraction might explain why users make more efficient decisions with abstract maps than with realistic displays under time pressure. In the long run, however, map-based decision makers may benefit from using a novel spatial display or map interaction tool: If decision-makers know how to efficiently apply novel displays and tools, the benefit-cost ratio of these displays and tools might outperform the benefit-cost ratio of familiar maps that are currently in use.

Generally, the task seems to dictate which map types are more suitable for making accurate decisions, even for a road selection task: On the one hand, participants are more accurate in selecting the fastest route in drive time with road maps. One plausible explanation is that road maps contain more task-relevant information, such as a visually salient classification of the road network. These results replicate findings of Gill (1993), who observed that people perform better at selecting the fastest route when the road classification is depicted using distinct and unambiguous line ordering and colors.

On the other hand, participants are more accurate in selecting the shortest routes in distance with satellite images, probably because they are distracted or misled by the base map information showing the road network. The road network information is not relevant for selecting the “shortest route” and can even be regarded as visual clutter in this task. The fact that the tested road maps on average are more cluttered (measured by Rosenholtz’ subband entropy) than the satellite images might serve as another explanation for the relatively poor accuracy of participants with road maps in the “shortest route” task.

While accuracy differences among the display types seem to depend on the concrete road selection task, it is striking that people are significantly more confident in satellite images**.** Possible explanations for this could be “naïve realism” or over-confidence in realistic displays.

These are phenomena for which other empirical studies have presented evidence (Fabrikant and Boughman, 2006; Hegarty et al., 2009; Smallman and St. John, 2005; Zanola et al., 2009). My experimental findings indicate that the results of these authors seem to be valid also for road selection tasks and under time pressure conditions.

When selecting slopes, participants perform better with slope maps and contour maps than with shaded relief maps, both in terms of accuracy and confidence. High confidence ratings suggest that people did not choose the correct locations randomly, but that the information on the slope maps has indeed supported people actively in their decision-making process. Confirming long-standing (but rarely empirically validated) cartographic design theory (Bertin, 1967), I find that people make better decisions with well-designed contour and slope maps. In contrast, more realistic displays do not seem to enhance efficient and effective map-based decision making in a slope detection context, which is in accordance with early empirical studies on hillshading maps (Potash et al., 1978). Although shaded relief maps contain more information than the contour lines maps, both participants' accuracy and confidence are worse with shaded relief maps. Seemingly, the shaded relief information confuses participants more than it facilitates obtaining slope information. The poor performance with shaded relief maps is especially striking under short time limits. This is congruent with participants' qualitative answers from Experiment I, in which relief information was regarded as aesthetic detail, which is not helpful for making efficient decisions under time pressure. The low confidence in relief maps in slope detection tasks is in concert with low preference and performance results for relief maps in previous work in a wayfinding experiment (Soh and Smith-Jackson, 2004). Low overall accuracy and confidence measures for relief maps indicate that representing elevation in an implicit way – which makes the spatial display more cluttered – is not an efficient method for slope detection tasks under time pressure. One reason for this might be that processing the implicit information on these hillshading maps might require a considerable amount of time.

The findings from the slope detection experiment suggest that the benefit of explicitly communicating thematically relevant information, even in a graphically abstract way (i.e., at a potentially higher cognitive cost), is greater for efficient and effective map-based decision making, than adding preferred and attractive, but visually more cluttered realism (i.e., higher perceptual cost). Low participant performance with shaded relief maps – even lower than with more abstract contour maps, containing even less information – indicates that visual realism might negatively influence map-based decision making, especially in time-critical contexts. In sum, realistic maps do not seem to support *efficient* (i.e., accurate **and** fast) decision making, because participants seem to have problems making *effective* (i.e., accurate) decisions under time pressure with these maps.

9.3 Interactivity

Previous work on interactivity has claimed that the 2D tools zooming and panning are the two most important tools for interactive maps and more important than the 3D tools rotating and tilting (Harrower and Sheesley, 2005). The experimental results in this study confirm this view: Firstly, the 2D interaction tools zooming and panning are regarded as the most suitable interaction tools for road selection tasks both under time pressure (TP) and without time pressure (NTP). The 3D interaction tools, rotating and tilting, seem less important. While the rotation tool seems to be the least important interaction tool both under and without time pressure, the tilting tool is regarded as significantly less suitable under time pressure. In qualitative, open-ended answers, participants mentioned tilting and rotating are “fun to use”, and useful to explore the area. Overall, these experimental results suggest that the 2D interaction tools are more important especially in time-critical situations. Seemingly, the 3D interaction tools tilting and rotating are considered rather as “toys” than as particularly helpful tools for efficient map-based decision making. Secondly, people actually use 2D interaction tools more often than 3D tools when they solve different tasks with an interactive map. The experimental findings provide evidence for a robust, task-independent, pattern that participants actually zoom and pan more than they rotate and tilt.

While interacting with the display can be important to make effective decisions, interaction in a 3D display is often regarded as cognitive cost (Bleisch, 2011; Nielsen, 2007; Shepherd, 2008) and thus seems to impair efficient decision making. For instance, people often have difficulties in orientating when interacting with a 3D display (Nielsen, 2007). The decision to interact with the map can be regarded as a decision which is made based on the assumption that the benefit of the interaction will exceed the cognitive cost.

Taking this into consideration, the results suggest that users believe that panning and zooming have the highest benefit-cost ratio of all interaction tools. In contrast, the benefits for using the 3D interaction tools under time pressure do not seem to outweigh the costs (i.e., the effort and time spent to interact). 3D interaction tools (such as tilting) seem to be only suitable when the task is complex (such as detecting the steepest slope) and map users are not under time pressure.

When aggregating frequencies of interactions across all tasks, I find that people overall use each of the four interaction tools more when not under time pressure. This suggests that people might believe that the benefit-cost ratio of map interactions is higher without time pressure. However, comparing interaction frequencies across the tasks, these TP/NTP differences in tool usage vary a lot among the tasks: For instance, when detecting the steepest slope, users tilt significantly more when there are no temporal constraints. In contrast, participants rotate significantly more without time pressure when they have to determine the highest point of a path. Finally, TP/NTP differences for panning are only significant when participants are assessing the elevation of two given points. Overall, these results show that it is both tool- and task-dependent to what extent time pressure affects human-map interactions.

Furthermore, I did not find strong evidence that participants who interact more with the map would be more successful in map-based decision making tasks. This finding supports empirical results of other authors researching the benefits of interactivity (Keehner et al., 2008), and indicates that people who interact more with maps do not necessarily make more effective or efficient map-based decisions. In my experimental work, interacting with the map does positively influence response accuracy only on one occasion: Participants who zoom in and out more frequently are more accurate in the most complex of these tasks without time pressure. This suggests that the benefit-cost ratio of the zooming tool actually might be higher than the benefit-cost ratios of the other tools, especially in complex, time-consuming map use tasks.

In summary, the experimental results allow for the following conclusion about the effect of map-related factors on the quality of map-based decisions: ***The effect of interactivity on the efficiency and effectiveness on map-based decisions is considerably lower than the effect of different map types and map designs.***

9.4 User characteristics (Individual and group differences)

Differences in rotation abilities and sex influence participant preferences for map types, and ways to interact with the map. For example, high-spatial participants seem to have a higher preference to zoom, pan and rotate a map. This is in concert with findings by Cohen and Hegarty (2007), who found that participants with good internal visual abilities are more (and not less) likely to rotate external 3D visualizations, as the productive use of an external visualization depends on good internal visualization abilities. In contrast, low-spatial participants prefer to tilt spatial displays when solving road selection tasks more than high-spatial participants. Sex seems to influence the preference for map types, as female participants prefer – in some occasions significantly – road maps more strongly compared to males.

Previous map-related studies have found that males are more confident in map-related decisions than females (Furnham, 2001; Furnham et al., 1999; Lloyd et al., 2002). I could replicate this pattern in a road selection task under time pressure: Males are significantly more confident in their road selection choices than females, while females are (slightly) more accurate in this task. However, this sex difference in confidence is only significant for short time limits. In Experiment III, a slope detection experiment with a similar setup, female accuracy is slightly above male accuracy, and overall male confidence ratings are higher than female's. The male/female confidence differences are again larger for shorter response time limits. However, in Experiment III none of these differences are statistically significant. Thus, it is unclear if map-based decision making under time pressure can be regarded as another context in which males might overestimate and females underestimate their performance.

Overall, there seems to be a robust pattern implying that higher confidence of male participants is more evident under short time limits. This might indicate that males believe to be efficient (i.e., confident under short limits) map-based decision-makers, while females might be the more effective (i.e., accurate and confident, but not necessarily under short limits) ones. This is somewhat in accordance with previous studies reporting that females take longer to complete map-related tasks (Delvin and Bernstein, 1997; Galea and Kimura, 1993), and it suggests that time pressure might have a more negative effect on female map users than on males. One explanation for this is that there is a tendency that females take longer to visually process images (Bunch and Lloyd, 2006), as significant sex differences in mental rotation – measured under time pressure – also suggest.

Solving several tasks using a virtual globe, participants who play video games on a regular basis are significantly more confident in their answers. These results suggest that playing video games might not only have a positive effect on spatial abilities (Feng et al., 2007; Terlecki et al., 2008), but also on the confidence in answers. One interpretation of this is that people who play video games frequently might feel more familiar when solving interactive tasks in a 3D virtual environment. Participants' background and training (e.g., sex, map use experience, virtual globe usage, and rotation of paper maps) does not significantly affect people's accuracy and confidence or how they interact with virtual globes.

9.5 Limitations of the empirical work

In the remainder of this chapter, I survey to what extent my experiments have contributed to exploring the effect of time pressure on map-based decision making. The strengths, weaknesses, potential routes to overcoming the weaknesses, and the implications of the strengths for future work are listed below in Table 8.

Table 8: Overview of limitations of the four experiments

	Strengths	Weaknesses	Routes to overcoming weaknesses	Implications of strengths for future work
Exp I	- Assessment of preferences for interaction tools and map types under different contexts - Assessment of the effect of user-related factors on map preferences	- Preferences might be task-dependent - Map type ratings might be due to (good or poor) map design of sample maps - An identical response time limit is perceived differently by different participants	- Assessing preferences for different tasks for more robust results - Testing different map designs for one map type - Measuring individual stress levels	- Future work can focus on the most "successful" map types and interaction tools - Preferences can be compared with performance (accuracy and confidence)
Exp II	- Assessment of the effect of time pressure, task complexity and map types in a road selection task - Large number of road maps and satellite images	- Findings are limited to route selection in a 2D environment - Labeling of the road maps and satellite images might have influenced confidence	- Choosing alternative tasks (e.g., in a 3D environment) - Omitting labels in road maps (or adding labels in satellite images)	- Robustness of speed-accuracy and speed-confidence trade-offs might be validated with other tasks - Robustness of accuracy/confidence discrepancy between road maps and satellite images can be further investigated
Exp III	- Assessment of the effect of time pressure and different map types in a slope detection task - Application of signal detection theory (SDT) to different shaded relief maps	- Results (i.e., low confidence in realistic maps) might be due to participant sample consisting of expert users - Findings might be biased by difference in information content (i.e., contour line interval)	- Conducting the experiment with cartographic novices - Varying contour line intervals	- Maps should be designed where the thematically relevant information is salient (like in the slope maps) - An alternative pattern of a time pressure effect (i.e., inverted U-shaped curve) has been found in empirical work - Accuracy can be analyzed in more depth by SDT
Exp IV	- Four different map use tasks - Comparison of tool preferences and tool usage (human-map interaction) - Assessment of the effect of interaction on accuracy and confidence	- Lack of significant findings of user characteristics might be due to small sample size (N=21)	- Conducting the experiment with a larger sample	- Focus on panning and zooming - Better interactive visualizations have to be designed that actually can enhance efficient decision making

Table 8 shows that all of the experiments indeed have their strengths and limitations. However, each experiment has contributed to a better understanding of map-based decision making under time pressure. The implications of these experiments for new research avenues are demonstrated in the next chapter.

10. Conclusion

The aim of this thesis was to assess which role time pressure plays for map-based decisions. This endeavor was guided by the following overall research question:

What is the effect of time pressure on map-based decision making considering varying spatial display types and map designs, interaction tools, and users with varying backgrounds?

Four experiments and several expert interviews have been conducted to shed light on this question. In the remainder of this section, I will summarize the main findings of this thesis and answer the more specific research questions, as formulated in section 1.3. Thereby I will show how this thesis has contributed to bridging the research gap between time pressure research and empirical map design and map use studies. In the second part of this chapter, I will discuss possible directions for future work resulting from the limited scope of this thesis.

10.1 Summary

As for time pressure, I could show that short response time limits affect user preferences, response accuracy and response confidence for map-based decision making to a great extent. In a first experiment, different decision time limits affected the preference ratings for several map types and interaction tools. In follow-up experiments, I demonstrated that time pressure also affects response accuracy, and – to an even larger degree – also response confidence. In one experiment on slope detection, time pressure had a positive effect on map-based decision making: Both overall response accuracy and confidence followed an “inverted U-shaped curve” pattern. That is, participants were most accurate and confident at a moderate time limit, while their accuracy and confidence decreased with more decision time available. Finally, time pressure also influenced the frequency of human-map interactions in an experiment with a virtual globe, as participants interacted more with the virtual globe when they were not under time pressure.

People’s preferences for spatial display types and map designs seem to depend on the decision time for decision making: While road maps seem to be the most preferred display type for a road selection task, regardless of time pressure, participants seem to regard satellite images as more suitable for situations that do not involve time pressure.

Different spatial display types and map designs influenced the efficiency and effectiveness of map-based decisions as well: Firstly, in a road selection task, participants overall were more confident in more realistic satellite images than in road maps, while they were not more accurate with the satellite images. Secondly, in a slope detection task, participants were less accurate and less confident with realistic shaded relief maps than with contour and slope maps, especially under time pressure. Overall, my experimental results and expert interviews support the view that realistic representations are not particularly efficient under time pressure. As for interactivity, there is no strong evidence that users who interact with the map are more accurate or more confident in their map-based decisions.

Finally, user-related factors also influence map-based decisions to a certain degree. For instance, differences in rotation abilities and sex affect participant preferences for map types, and ways to interact with the map. While differences in rotation abilities seem to have a strong effect on interaction tool preferences, sex rather influences map display preferences. In general, males tend to have a higher confidence in their map-based decisions than females, while their response accuracy is not significantly higher than the response accuracy of females.

In summary, the empirical work has shown that time pressure indeed has an effect on map-based decision making, and strongly interacts with other map- and user-related factors. Time pressure seems to be an important factor for the effectiveness and efficiency of decisions made based on maps and geovisualisations, and should therefore be further investigated in cognitive and usability studies. With this finding in mind, a lot of further research avenues are emerging in the domain of empirical map use studies under time pressure, which I will discuss in the remainder of this chapter.

10.2 Outlook

This thesis has shed light on different issues concerning the efficiency and effectiveness of maps, which are relevant for the disciplines of geography, cartography, and cognate research fields like geovisualization. As mentioned in Chapter 1, it is a main interest of these disciplines to identify and produce maps that support efficient and effective spatio-temporal decision-making. The empirical results of this thesis imply that cartographers should focus on maps that depict the thematically relevant information in a perceptually salient way, especially for time pressure situations. For other map use contexts, details of the landscape, realistic satellite imagery, shaded reliefs, or panoramic views might be aesthetically pleasing. However, if these map elements are not relevant for the task at hand under time pressure, they should not be displayed in maps.

As mentioned in section 9.5, there are certain limitations to the empirical results of this thesis. For instance, within the scope of this thesis, the effect of time pressure on map-based decision making could only be investigated within a limited series of experiments, scenarios and tasks.

Future experiments should be conducted in different map-based decision contexts under time pressure, in order to further investigate the robustness and generalizability of the findings discussed already in section 9.5.

This thesis has shown that both speed-performance (i.e., speed-accuracy and speed-confidence) tradeoffs and inverted U-shaped curves can be found in map-based decision-making tasks under time pressure. This implies that it is very context- and task-dependent how time pressure influences spatial decision making. One key research avenue of future empirical work could focus on the question how particular map-based decision-making contexts influence the relationships between decision time limits and performance measures. In particular, I suggest that further research regarding the effect of time pressure on map-based decision making should focus on the following three hypotheses that emerge from my empirical work:

1. The speed-confidence trade-off is stronger than the speed-accuracy trade-off.
2. Both of these trade-offs are more striking between short and moderate (e.g., the average time pilot testers need) time limits than between moderate and generous time limits.
3. Certain map-based decisions follow an inverted U-shaped curve (i.e., accuracy and confidence tend to decrease with more response time).

As mentioned above, in this study time pressure was operationalized using identical time *limits* for each participant. Futures studies could also focus on the concept of time *pressure* by actually measuring whether participants experienced or felt any pressure when making map-based decisions. One way to do this would be to ask participants whether they felt any stress while performing tasks. Another – probably more complex and time-consuming – approach could involve measuring individual stress levels or employing fMRI (functional magnetic resonance imaging) methods, that is, measuring changes in blood flow related to the neural activity in the brain. This might provide direct insights on cognitive processes in human brains during making map-based decisions.

Another key research avenue could be studying the effect of certain map-related factors, and how they interact with context-related factors, such as, task complexity or time pressure. In general, this study has shown that realistic 3D maps are not as efficient (for typical map-based decisions under time pressure in road selection and slope detection tasks) as one might think. Future work should identify whether there are tasks where realistic displays actually can enhance efficient map-based decision making, in order to investigate the usefulness of these visualizations for time-critical situations. Moreover, very little is still known about the effect of time pressure on map-based decision making with other cartographic representations, such as thematic maps or cartograms. However, there might be certain map use tasks where these representations are more efficient and effective than the maps used in this study.

Future map use studies could also focus on the effect of map design beyond that it has been possible within the scope of this thesis. For instance, one could compare road maps or slope maps with different color schemes, levels of complexity or visual clutter, in order to assess what might be the "optimal level" of complexity or clutter for efficient map-based decision making. Cartographers could take these "optimal levels" into account when designing maps for time-critical decision making settings.

Furthermore, this thesis focused on spatial displays on large computer screens. Hence, display size was one factor which was held constant throughout all experiments. As map applications on mobile devices with small displays, such as PDAs and smartphones, keep growing in popularity, it seems important to evaluate how users perform under time pressure with those mobile devices in the field. Therefore, I suggest validating the results of this thesis by conducting similar experiments on standard mobile devices as another future key research avenue for geovisualization.

As for user-related factors, it is important to note that my test samples chiefly consisted of a homogenous set of users (geography and cartography students) with similar background and training. It is unclear at this point how performance is affected by different degrees of map-use experience and how user background interacts with map-related factors. In future work, map-reading novices should be tested in similar time pressure contexts, in order to compare findings of previous studies by Hegarty and colleagues (2009), who have found higher preferences for 3D maps among "naïve cartographers". Future studies could focus on the Dunning-Kruger effect, which suggests that unskilled people tend to overestimate their own performance, and that highly skilled persons tend to underrate their abilities (Kruger and Dunning, 1999).

Moreover, further research on high-spatial and low-spatial participants could investigate the question how these categories influence user performance with more or less abstract maps, similar to work by Dillemuth (2005), who has shown that high-spatial participants seem to perform better with generalized maps than participants with less spatial-visual skills. Such studies could inform design guidelines of maps for certain users. Furthermore, it might be interesting how alternative group effects, such as the 2D/4D ratio (i.e., the ratio of the lengths of the index and ring finger), right-handedness (Lloyd and Bunch, 2005), or cultural differences might interact with map-related or context-related factors and the dependent variables accuracy and confidence. As previous work has revealed that cultural differences predict decision making accuracy and response time in a wayfinding task (Soh and Smith-Jackson, 2004), the cultural background of the decision-maker might also influence the quality of decisions in other map use tasks.

To better understand map-based decision making under time pressure, it might also be interesting to study other expert groups in this field, such as, soldiers, orienteers, or sailors. Studying them might contribute to additional insights about typical domains of map-based decision making under time pressure, such as military, orienteering or sailing. As these people make map-based decisions under time pressure in their everyday lives, the knowledge of these people might help improving maps and other spatial displays for efficient and effective decision making.

Finally, I also encourage like minded researchers in GIScience and cartography to more often try to analyze response accuracy with the signal detection approach. Future empirical map design and map use studies could focus more on the question of what kinds of errors (false alarms or misses) might result in low accuracy rates, and also take into account various signal detection measures such as hit rates, false alarm rates, discriminability and bias. This might in turn lead to more focused map design guidelines for decision making in general, and under time pressure in particular.

References

Abend, P., Thielmann, T., Ewerth, R., Seiler, D., Mühling, M., Döring, J., Grauer, M., and Freisleben, B. (2011). Geobrowsing the globe: A geovisual analysis of Google Earth usage. In, *Proceedings of GeoViz 2011*, Hamburg.

Ahituv, N., Igbaria, M., and Sella, A. (1998). The effects of time pressure and completeness of information on decision making. *Journal of Management Information Systems,* vol. 15, no. 2: 153-172.

Albert, W. S. and Golledge, R. G. (1999). The use of spatial cognitive abilities in geographical information systems: The map overlay operation. *Transactions in GIS,* vol. 3 no. 1: 7-21.

Andrews, F. M. and Farris, G. F. (1972). Time pressure and performance of scientists and engineers: A five-year panel study. *Organizational Behavior and Human Performance,* vol. 8, no. 2: 185-200.

Andrienko, N., Andrienko, G., Voss, H., Bernardo, F., Hipolito, J., and Kretchmer, U. (2002). Testing the usability of interactive maps in CommonGIS. *Cartography and Geographic Information Science,* vol. 29, no. 4: 325-342.

Aretz, A. J. and Wickens, C. D. (1992). The mental rotation of map displays. *Human Performance,* vol. 5, no. 4: 303-328.

Bailenson, J. N., Shum, M. S., and Uttal, D. H. (1998). Road climbing: Principles governing asymmetric route choices on maps. *Journal of Environmental Psychology,* vol. 18: 251-264.

Barkley, R. A. (1991). The ecological validity of laboratory and analogue assessment methods of ADHD symptoms. *Journal of Abnormal Child Psychology,* vol. 19, no. 2: 149-178.

Baus, J., Krüger, A., and Wahlster, W. (2002). A resource-adaptive mobile navigation system. In, *Proceedings of the 7th International Conference on Intelligent User Interfaces*, ACM New York, New York, USA: 15-22.

Bentham, J. (1789). *An introduction to the principle of morals and legislations*, Blackwell, Oxford.

Bertin, J. (1967). *Sémiologie graphique*, Mouton, Paris.

Birenbaum, M., Kelly, A., and Levi-Keren, M. (1994). Stimulus features and sex differences in mental rotation test performance. *Intelligence,* vol. 19: 51-64.

Bleisch, S. (2011). *Evaluating the appropriateness of visually combining quantitative data representations with 3D desktop virtual environments using mixed methods*, Doctoral thesis at the City University of London.

Bloom, G. S. (2007). *Helicopters: How They Work.* (http://www.helicopterpage.com/html/forces.html)

Board, C. (1978). Map reading tasks appropriate in experimental studies in cartographic communication. *The Canadian Cartographer,* vol. 15: 1 - 12.

Boughman, A. (2005). *Realism and connotations of data quality in computer-displayed maps*, Unpublished master thesis at the University of California, Santa Barbara.

Brophy, D. M. (1980). Some reflections on the complexity of maps. *Technical papers of ACSM 40th Annual Meeting, St. Louis*: 343-352.

Brunyé, T. T., Mahoney, C. R., and Gardony, A. L. (2010). North is up(hill): Route planning heuristics in real-world environments. *Memory & Cognition,* vol. 38, no. 6: 700-712.

Bunch, R. L. and Lloyd, R. E. (2006). The cognitive load of geographic information. *The Professional Geographer,* vol. 58: 209-220.

Busemeyer, J. R. and Townsend, J. T. (1993). Decision-field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review,* vol. 100, no. 3: 432-459.

Canham, M., Smallman, H., and Hegarty, M. (2007). Using complex visual displays: When users want more than is good for them. In: Mosier, K., and Fischer, U., (eds.), *Proceedings of the Eighth International NDM Conference*, Pacific Grove, CA.

Cartwright, W., Crampton, J., Gartner, G. F., Miller, S., Mitchell, K., Siekierska, E., and Wood, J. (2001). Geospatial information visualization user interface issues. *Cartography and GIS* vol. 28, no. 1: 45-60.

Chang, K. S., Antes, J., and Lenzen, T. (1985). The effect of experience on reading topographic relief information: Analyses of performance and eye movements. *The Cartographic Journal,* vol. 22: 88-94.

Choi, J. S., I (2003). Processes underlying sex differences in route-learning strategies in children and adolescents. *Personality and Individual Differences,* vol. 34: 1153-1166.

Christenfeld, N. (1995). Choices from identical options. *Psychological Science,* vol. 6: 50-55.

Cohen, C. A. and Hegarty, M. (2007). Individual differences in use of external visualisations to perform an internal visualisation task. *Applied Cognitive Psychology,* vol. 21: 701-711.

Collier, P., Forrest, D., and Pearson, A. (2003). The representation of topographic information on mpas: The depiction of relief. *The Cartographic Journal,* vol. 40: 17-26.

Coors, V., Elting, C., Kray, C., and Laakso, K. (2005). Presenting route instructions on mobile devices: From textual directions to 3D visualization. In: Dykes, J., Kraak, M.J., and MacEachren, A.M., (eds.), *Exploring Visualization*, Elsevier, Oxford: 529-550.

Crampton, J. W. (2002). Interactivity types for geographic visualization. *Cartography and Geographic Information Science,* vol. 29, no. 2: 85-98.

Dahlke, J. and Winck, D. (2009). 3D worlds for the Bundeswehr. *Proceedings*, Civil Commercial Imagery Evaluation Workshop, Fairfax Oaks, VA, April 2, 2009.

Delvin, A. S. and Bernstein, J. (1997). Interactive way-finding: Map style and effectiveness. *Journal of Experimental Psychology,* vol. 17: 99-110.

Dillemuth, J. (2005). Map-design evaluation for mobile displays. *Cartography and Geographic Information Science,* vol. 32, no. 4: 285-301.

Dillemuth, J. (2009). Navigation tasks with small-display maps: The sum of the parts does not equal the whole. *Cartographica,* vol. 44, no. 3: 187-200.

Dobson, M. W. (1980). The influence of the amount of graphic information on visual matching. *The Cartographic Journal,* vol. 17, no. 1: 26-32.

Dobson, M. W. (1983). Visual information processing and cartographic communication: The utility of redundant stimulus dimensions. In: Taylor, D.R.F., (ed.), *Graphic communication and design in contemporary cartography* vol. 2, John Wiley & Sons, Chichester: 149-175.

Downing, R., Moore, J. L., and Brown, S. W. (2005). The effects and interaction of spatial visualization and domain expertise on information seeking. *Computers in Human Behavior* vol. 21: 195-2009.

Duckham, M. and Kulik, L. (2003). "Simplest" paths: Automated route selection for navigation. In: Kuhn, W., Worboys, M.F., and Timpf, S., (eds.), *Spatial Information Theory. 6th International Conference, COSIT 2003, Ittingen, Switzerland,* vol. 2825, Springer Berlin, Heidelberg: 169-185.

Dykes, J., MacEachren, A., and Kraak, M.-J. (2005). Exploring geovisualization. In: Dykes, J., MacEachren, A.M., and Kraak, M.J., (eds.), *Exploring geovisualization*, Elsevier, Amsterdam: 3-19.

Eckert, M. (1921). *Die Kartenwissenschaft: Forschungen und Grundlagen zu einer Kartographie als Wissenschaft*, W. de Gruyter, Berlin.

Edland, A. and Svenson, O. (1993). Judgment and decision making under time pressure: Studies and findings. In: Svenson, O., and Maule, J., (eds.), *Time pressure and stress in human judgment and decision making*, Plenum Press, New York, London: 27-40.

EWZ (2010). *Energy moving with the times. ewz - the company.* (http://www.stadt-zuerich.ch/content/dam/stzh/ewz/Deutsch/Ueber%20ewz/Publikationen%20und%20Broschueren/Unternehmensbroschuere_e.pdf)

Fabrikant, S. I. (2005). Towards an understanding of geovisualization with dynamic displays: Issues and prospects. In, *Proceedings of the American Association for Artificial Intelligence (AAAI) 2005 Spring Symposium Series: Reasoning with mental and external diagrams*, Stanford, CA: 6-11.

Fabrikant, S. I. and Goldsberry, K. (2005). Thematic relevance and perceptual salience of dynamic geovisualization displays. In, *Proceedings of the ICA/ACI International Cartographic Conference*, A Coruña, Spain, Jul. 9-16, 2005.

Fabrikant, S. I. and Boughman, T. (2006). Communicating data quality through realism. In, *Proceedings of GIScience 2006*, Münster, Germany: 59-60.

Fabrikant, S. I. and Lobben, A. (2009). Special issue on cognitive issues in geovisualization. *Cartographica,* vol. 44, no. 3.

Fabrikant, S. I., Rebich-Hespanha, H., and Hegarty, M. (2010). Cognitively inspired and perceptually salient graphic displays for efficient spatial inference making. *Annals of the Association of American Geographers,* vol. 100, no. 1: 13-29.

Fabrikant, S. I., Rebich-Hespanha, H., Andrienko, N., Andrienko, G., and Montello, D. R. (2008). Novel method to measure inference affordance in static small-multiple map displays representing dynamic processes. *The Cartographic Journal,* vol. 45, no. 3: 201-215.

Fairbairn, D. (2006). Measuring map complexity. *The Cartographic Journal,* vol. 43, no. 3: 224-238.

Feng, J., Spence, I., and Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychological Science,* vol. 8: 850-855.

Fields, A. (2009). *Discovering statistics using SPSS. Third edition*, Sage, London.

Förster, J., Higgins, E. T., and Taylor Bianco, A. (2003). Speed/accuracy decisions in task performance: Built-in trade-off or separate strategic concerns? *Organizational Behaviour and Human Decision Processes,* vol. 90: 148-164.

Friedman, M. and Savage, L. J. (1948). The utility analysis of choices involving risk. *The Journal of Political Economy,* vol. 56, no. 4: 279-304.

Furnham, A. (2001). Self-estimates of intelligence: Culture and gender differences in self and other estimates of both general (g) and multiple intelligences. *Personality and Individual Differences,* vol. 31: 1381-1405.

Furnham, A., Fong, S., and Martin, N. (1999). Sex and cross-cultural differences in the estimated multi-faceted intelligence quotient score for self, parents, and siblings. *Personality and Individual Differences,* vol. 26: 1025-1034.

Gahegan, M. (2008). Multivariate geovisualization. In: Wilson, J.P., and Fotheringham, A.S., (eds.), *The Handbook of Geographic Information Science*, Blackwell Publishing: 292-316.

Galea, L. A. and Kimura, D. (1993). Sex differences in route learning. *Personality and Individual Differences,* vol. 14: 53-65.

Garlandini, S. and Fabrikant, S. I. (2009). Evaluating the effectiveness and efficiency of visual variables for geographic information visualization. In: Stewart Hornsby, K., Claramunt, C., Denis, M., and Ligozat, G., (eds.), *Spatial Information Theory. 9th International Conference, COSIT 2009, Aber Wrac'h, France.,* vol. 5756, Lecture Notes in Computer Science, Springer, Berlin, Heidelberg: 195-211.

Garrett, H. E. (1922). A study of the relation of accuracy to speed. *Archives of Psychology*, no. 56.

Gigerenzer, G. (2002). The adaptive toolbox. In: Gigerenzer, G., and Selten, R., (eds.), *Bounded rationality: The adaptive toolbox*, Dahlem Workshop Reports: 37-50.

Gill, G. (1993). Road map design and route selection. *The Cartographic Journal,* vol. 30: 163-166.

Goldin, S. E. and Thorndyke, P. W. (1981). *An Analysis of Cognitive Mapping Skill,* Rand, Santa Monica.

Golledge, R. G. (1995). *Path selection and route preference in human navigation: A progress report. Working paper UCTC No. 277,* The University of California, Transportation Center, Berkeley.

Griffin, A. and Bell, S. (2009). Applications of signal detection theory to geographic information science. *Cartographica,* vol. 44, no. 3: 145-158.

Harrower, M. and Sheesley, B. (2005). Designing better map interfaces: A framework for panning and zooming. *Transactions in GIS,* vol. 9, no. 2: 77-89.

Harrower, M., MacEachren, A. M., and Griffin, A. L. (2000). Developing a geographic visualization tool to support Earth science learning. *Cartography and Geographic Information Science,* vol. 27, no. 4: 279-293.

Hegarty, M. (2010). Components of spatial intelligence. In: Ross, B.H., (ed.), *The psychology of learning and motivation,* vol. 52, Academic Press, San Diego: 265-297.

Hegarty, M., Smallman, H. S., and Stull, A. T. (2008). Decoupling of Intuitions and Performance in the Use of Complex Visual Displays. In: Love, B., McRae, K., and Sloutsky, V., (eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, Washington, DC: 881-886.

Hegarty, M., Canham, M. S., and Fabrikant, S. I. (2010). Thinking about the weather: How display salience and knowledge affect performance in a graphic interference task. *Journal of Experimental Psychology: Learning, Memory and Cognition,* vol. 36, no. 1: 37-53.

Hegarty, M., Smallman, H. S., Stull, A. T., and Canham, M. (2009). Naive geography: How intuitions about display configuration can hurt performance. *Cartographica,* vol. 44, no. 3: 171-186.

Hegarty, M., Keehner, M., Cohen, C. A., Montello, D. R., and Lippa, Y. (2007). The role of spatial cognition in medicine: Applications for selecting and training professionals. In: Allen, G., (ed.), *Applied spatial cognition*, Lawrence Erlbaum Associates, Mahwah, NJ.

Hochmair, H. (2009). The influence of map design on route choice from public transportation maps in urban areas. *The Cartographic Journal,* vol. 46, no. 3: 242-256.

Hochmair, H. and Rinner, C. (2005). Investigating the need for eliminatory constraints in the user interface of bicycle route planners. In: Cohn, A.G., and Mark, D.M., (eds.), *Spatial Information Theory. 7th International Conference, COSIT 2005, Ellicottville, NY,* Springer, Berlin, Heidelberg: 49-64.

Hwang, M. I. (1994). Decision making under time pressure: a model for information systems research. *Information and Management,* vol. 27: 197-203.

International Standards Organisation, I. (1998). ISO 2941-11 Economic requirements for office work with visual display terminals (VDT)s - Part 11. *Guidance on Usability, Geneve*: 22.

Ishikawa, T. and Yamazaki, T. (2009). Showing where to go by maps or pictures: An empirical case study at subway exits. In: Stewart Hornsby, K., Claramunt, C., Denis, M., and Ligozat, G., (eds.), *Spatial Information Theory. 9th International Conference, COSIT 2009, Aber Wrac'h, France.,* vol. 5756, Springer, Berlin, Heidelberg: 330-341.

Itti, L. and Koch, C. (2001). Computational Modelling of Visual Attention. *Nature Reviews Neuroscience* vol. 2, no. 3: 194-203.

James, K. H., Humphrey, G. K., Vilis, T., Corrie, B., Baddour, R., and Goodale, M. A. (2002). "Active" and "passive" learning of three-dimensional object structure with an immersive virtual environment. *Behaviour Research Methods, Instruments & Computers,* vol. 34: 383-390.

Jarvis, A. R., HI, Nelson, A., and Guevara, E., 2008, Hole-filled seamless SRMT data V4: International Centre for Tropical Agriculture (CIAT). .

Johnson, E. J., Payne, J. W., and Bettman, J. R. (1993). Adapting to time constraints. In: Svenson, O., and Maule, J., (eds.), *Time pressure and stress in human judgment and decision making*, Plenum Press, New York, London: 103-1116.

Jordan, K., Wüstenberg, T., Heinze, H., Peters, M., and Jäncke, L. (2002). Women and men exhibit different cortical activation patterns during mental rotation tasks. *Neuropsychologia* vol. 40: 2397-2408.

Kahneman, D. and Tversky, A. (2000). Choices, values and frames. In: Kahneman, D., and Tversky, A., (eds.), *Choices, values and frames*, Cambridge University Press, Cambridge: 1-16.

Keehner, M., Hegarty, M., Cohen, C., Khooshabeh, P., and Montello, D. R. (2008). Spatial reasoning with external visualizations: What matters is what you see, not whether you interact. *Cognitive Science,* vol. 32: 1099-1132.

Kimerling, A. J., Muehrcke, P. C., and Muehrcke, J. (2005). *Map use. Reading, analysis and interpretation. Fifth edition*, JP Publications, Madison, Wisconsin

Kirschenbauer, S. (2005). Applying "true 3D" techniques to geovisualisation: An empirical study. In: Dykes, J., MacEachren, A.M., and Kraak, M.J., (eds.), *Exploring Geovisualisation*, Elsevier, Amsterdam: 363–387.

Klippel, A., Freksa, C., and Winter, S. (2006). You-are-here maps in emergencies: The danger of getting lost. *Journal of Spatial Science,* vol. 51, no. 1: 117-131.

Kluckhohn, C. and Murray, H. A. (1953). *Personality, its nature, society, and culture*, Alfred Knopf, New York.

Kraak, M. J. (1993). Three-dimensional map design. *The Cartographic Journal,* vol. 30, no. 2: 188-194.

Kruger, J. and Dunning, J. (1999). Unskilled and Unaware of It: How Difficulties In Recognising One's Own Incompetence Lead to Inflated Self-Assessments. *Journal of Personality and Social Psychology,* vol. 77: 1121-1134.

Krygier, J. B., Reeves, C., DiBiase, D., and Cupp, J. (1997). Design, implementation and evaluation of multimedia resources for geography and earth science education *Journal of Geography in Higher Education,* vol. 21: 17-39.

Kuoa, E. L., MacEachren, A. M., and Kraak, M. J. (2006). Evaluating the usability of visualization methods in an exploratory geovisualization environment. *International Journal of Geographic Information Science,* vol. 20: 425-448.

Lang, N. C. (2010). *Toponyme in der Einsatzdisposition*, Unpublished master thesis at the Department of Geography, University of Zurich.

Linn, M. C. and Petersen, A. C. (1985). Emergence and characterization of gender differences in spatial abilities: A meta-analysis. *Child Development,* vol. 56: 1479-1498.

Lloyd, R. E. (1997). Visual search processes used in map reading. *Cartography* vol. 34, no. 1: 11-32.

Lloyd, R. E. and Bunch, R. L. (2003). Technology and map-learning: Users, methods and symbols. *Annals of the Association of American Geographers,* vol. 93, no. 4: 828-850.

Lloyd, R. E. and Bunch, R. L. (2005). Individual differences in map reading spatial abilities. Using perceptual and memory processes. *Cartography and Geographic Information Science,* vol. 32, no. 1: 33-46.

Lloyd, R. E. and Bunch, R. L. (2008). Explaining map-reading performance efficiency: Gender, memory and geographic information. *Cartography and Geographic Information Science,* vol. 35, no. 3: 171-202.

Lloyd, R. E. and Bunch, R. L. (2010). Learning geographic information from a map and text: Learning environment and individual differences. *Cartographica,* vol. 45, no. 3: 171-172.

Lloyd, R. E., Hodgson, M., and Stokes, M. (2002). Visual categorization with aerial photographs. *Annals of the Association of American Geographers,* vol. 92: 241-266.

Lobben, A. K. (2004). Tasks, strategies and cognitive processes associated with navigational map reading: A review perspective. *The Professional Geographer,* vol. 56: 270-281.

Lobben, A. K. (2007). Navigational map reading: Predicting performance and identifying relative influence of map-related abilities. *Annals of the Association of American Geographers,* vol. 97, no. 1: 64-85.

MacEachren, A. and Kraak, M.-J. (2001). Research challenges in geovisualization. *Cartography and Geographic Information Science,* vol. 28, no. 1: 3-12.

MacEachren, A. M. (1995). *How maps work: Representation, visualization and design*, The Guilford Press, New York.

MacEachren, A. M., Boscoe, F. P., Haug, D., and Pickle, L. W. (1998). Geographic visualization: Designing manipulable maps for exploring temporally varying georeferenced statistics. *Proceedings of the 24th IEEE Symposium on Information Visualization (InfoVis '98), Los Alamitos, California, October 19-20, 1998*: 87-94.

March, J. G. (1994). *A primer on decision-making*, The Free Press, New York.

Marchak, F. M. and Zulager, D. D. (1992). The effectiveness of dynamic graphics in revealing structure in multivariate data. *Behaviour Research Methods, Instruments & Computers,* vol. 24: 253-257.

Martin, D. W. (2008). *Doing psychology experiments*, vol. 7, Thomson, Belmont, CA.

Maule, A. J. (1998). Strategies for coping with time pressure. In: Flin, R., Salas, E., Strub, M., and Martin, L., (eds.), *Decision making under stress*, Ashgate, Aldershot: 271-293.

Maule, A. J. and Svenson, O. (1993). Approaches to behavioral decision making. In: Svenson, O., and Maule, J., (eds.), *Time pressure and stress in human judgment and decision making*, Plenum Press, New York, London: 3-25.

Maule, A. J. and Edland, A. (1997). The effects of time pressure on judgment and decision making. In: Ranyard, R., Crozier, R., and Svenson, O., (eds.), *Decision making: Cognitive models and explanations*, Routledge, London: 198-204.

Maule, A. J. and Andrade, I. (1997). The effects of time pressure on decision making: How harassed managers cope. *Proceedings*, IEE Colloquium on Decision Making and Problem Solving.

Maule, A. J., Hockey, G. R. J., and Bdzola, L. (2000). Effects of time-pressure on decision-making under uncertainty: changes in affective state and information processing strategy. *Acta Psychologica,* vol. 104: 283-301.

Miller, J. G. (1960). Information input overload and psychopathology. *American Journal of Psychiatry,* vol. 116: 695-704.

Moellering, H. (1980). The real-time animation of three dimensional maps. *The American Cartographer,* vol. 7: 67-75.

Moellering, H. (2007). Expanding the ICA conceptual definition of a map. *Proceedings of the 23th International Cartographic Conference, Moscow.*

Montello, D. R. (2002). Cognitive map-design research in the twentieth century: Theoretical and empirical approaches. *Cartography and Geographic Information Science,* vol. 29, no. 3: 283-304.

Montello, D. R. and Freundschuh, S. M. (2005). Cognition of geographic information. In: B, M.R., and L, U.E., (eds.), *A research agenda for geographic information science*, CRC Press, Boca Raton, FL: 61-91.

Nielsen, A. (2007). *A qualification of 3D geovisualization*, Doctoral thesis at Aalborg University.

Ordoñez, L. and Benson III, L. (1997). Decisions under time pressure: How time constraints affects risky decision making. *Organizational Behaviour and Human Decision Processes,* vol. 71, no. 2: 121-140.

Ortner, F. (2011). *Dienstbasierte Kartengenerierung für Web-Atlanten unter Anwendung erweiterter OGC-Standards*, Unpublished master thesis at the Department of Geography, University of Zurich.

Palmer, J. (1994). Measuring the effect of attention on simple visual search. *Journal of Experimental Psychology: Human Perception and Performance,* vol. 19: 108-130.

Pareto, V. (1906). *Manuale di economia politica*, Edizione Critica, Milan.

Parsons, T. D., Larson, P., Kratz, K., Thiebaux, M., Bluestein, B., Buckwalter, J. G., and Rizzo, A. A. (2004). Sex differences in mental rotation and spatial rotation in a virtual environment. *Neuropsychologia 42,* vol. 42: 555-562.

Payne, J. (1982). Contingent decision behavior. *Psychological Bulletin,* vol. 93: 382-402.

Pedersen, D. M. (1999). Dimensions of environment competence. *Journal of Environmental Psychology,* vol. 19: 303-308.

Peruch, P., Vercher, J. L., and Gauthier, G. M. (1995). Acquisition of spatial knowledge through visual exploration of simulated environments. *Ecological Psychology,* vol. 7: 1-20.

Peters, L. H. and O'Connor, E. J. (1980). Situational constraints and work outcomes: The influences of a frequently overlooked construct. *Academy of Management Journal,* vol. 5: 391-397.

Pew, R. W. (1969). The speed-accuracy operating characteristics. In: Koster, W.G., (ed.), *Acta Psychologica 30: Attention and Performance II* vol. , North-Holland Publishing Company, Amsterdam: 16-26.

Philips, R. J., De Lucia, A., and Skelton, N. (1975). Some objective tests on the legibility of relief maps. *The Cartographic Journal,* vol. 12, no. 1: 39-46.

Pingel, T. J. (2010). *Strategic elements of route choice*, Dissertation at the University of California, Santa Barbara.

Pingel, T. J. and Clarke, K. C. (2005). Assessing the usability of a wearable computer system for outdoor pedestrian navigation. In, *Proceedings of Auto-Carto 2005, Las Vegas, Nevada*.

Potash, L. M., Farrell, J. P., and Jeffrey, T. (1978). A technique for assessing map relief legibility. *The Cartographic Journal,* vol. 15, no. 1: 28-35.

Prinzel, L. J. and Freeman, F. G. (1995). Sex differences in visuo-spatial ability: Task difficulty, speed accuracy tradeoff, and other performance factors. *Canadian Journal of Experimental Psychology,* vol. 49, no. 4: 530-539.

Rabin, M. (2000). Diminishing marginal utility of wealth cannot explain risk aversion. In: Kahneman, D., and Tversky, A., (eds.), *Choices, values and frames*, Cambridge University Press, Cambridge: 202-208.

Rega (2011). *Mission Statement.* (http://www.rega.ch/pdf/ueber_uns/Leitbild_en.pdf)

Riedl, A. (2006). Digital globes. In: Cartwright, W., Peterson, M.P., and Gartner, G.F., (eds.), *Multimedia Cartography (2nd ed),* Springer, Berlin: 255-266.

Robinson, A. H., Morrison, J. L., Muehrcke, P. A., Kimerling, A. J., and Guptill, S. (1995). *Elements of cartography,* vol. 6, John Wiley & Sons, New York.

Rosenholtz, R. (2001). Visual search for orientation among heterogeneous distractors: Experimental results and implications for signal-detection theory models of search. *Journal of Experimental Psychology: Human Perception and Performance,* vol. 27: 985-999.

Rosenholtz, R., Li, Y., and Nakano, L. (2007). Measuring visual clutter. *Journal of Vision,* vol. 7, no. 2: 1-22.

Rosenholtz, R., Li, Y., Mansfield, J., and Jin, Z. (2005). Feature congestion: A measure of display clutter. In, *Computer Human Interaction 2005*, ACM Press, Portland OR: 761-770.

Schöning, J., Hecht, B., Raubal, M., Krüger, A., Marsh, M., and Robs, M. (2008). Improving interaction with virtual globes through spatial thinking: Helping users ask "why?". In, *Proceedings of the 13th International Confererence on Intelligent User Interfaces*, Gran Canaria, Spain: 129-138.

Schultz, R. B., Kerski, J. J., and Pattersson, T. C. (2008). The use of virtual globes as a spatial teaching tool with suggestions for metadata standards. *Journal of Geography* vol. 107: 27-34.

Schutz & Rettung Zürich (2010a). *Protection & Rescue: Brochure.* (http://www.stadt-zuerich.ch/content/dam/stzh/pd/Deutsch/Schutz%20und%20Rettung/Ueber%20uns/Publikationen%20und%20Broschueren/Imagebroschuere_eng.pdf)

Schutz & Rettung Zürich (2010b). *Protection & Rescue "On call round the clock for you".* (http://www.stadt-zuerich.ch/content/dam/stzh/pd/Deutsch/Schutz%20und%20Rettung/Ueber%20uns/Publikationen%20und%20Broschueren/Imagebroschuere_eng.pdf)

Shepherd, I. D. H. (2008). Travails in the third dimension: A critical evaluation of three-dimensional geographical visualization. In: Dodge, M., McDerby, M., and Turner, M., (eds.), *Geographic Visualization: Concepts, Tools and Applications,* Wiley, Chichester: 199-222.

Sholl, M. J. and Liben, L. S. (1995). Illusory tilt and Euclidian schemes. *Journal of Experimental Psychology: Learning, Memory and Cognition,* vol. 21, no. 6: 1624-1638.

Simon, H. (1959). Theories of decision making in economics and behavioural science. *American Economic Review,* vol. 49, no. 3: 253-283.

Simon, H. A. and Larkin, J. H. (1987). Why a diagram is (sometimes) worth then thousand words. *Cognitive Science,* vol. 11: 65-100.

Slocum, T. A., Blok, C., Jiang, B., Koussoulakou, A., Montello, D. R., Fuhrmann, S., and Hedley, N. R. (2001). Cognitive and usability issuses in geovisualization. *Cartography and Geographic Information Science,* vol. 28, no. 1: 61-75.

Smallman, H. S. and St. John, M. (2005). Naive realism: Misplaced faith in realistic displays. *Ergonomics in Design,* vol. 13, no. 3: 6-13.

Smallman, H. S., St. John, M., Oonk, H. M., and Cowen, M. B. (2001). Information availabity in 2D and 3D displays. *Applied Perception,* vol. 21, no. 5: 51-57.

Smith, J. F., Mitchell, T. R., and Beach, L. R. (1982). A cost-benefit mechanism for selecting problem-solving strategies: Some extensions and empirical tests. *Organizational Behaviour and Human Performance,* vol. 29, no. 3: 370-396.

Soh, B. K. and Smith-Jackson, T. L. (2004). Influence of map design, individual differences, and environmental cues on wayfinding performance. *Spatial Cognition & Computation,* vol. 4, no. 2: 137-165.

Srinivas, S. and Hirtle, S. C. (2010). The role of motivation and complexity on wayfinding performance. In: Purves, R., and Weibel, R., (eds.), *GIScience 2010 Extended Abstracts Volume*, Zürich.

Stumpf, H. (1998). Gender-related differences in academically talented student's test scores and use of time on tests of spatial abilities. *Gifted Child Quarterly,* vol. 42: 157-171.

Suter, C. (2007). *Lawinenausbildung mit mobilen Systemen*, Unpublished diploma thesis at the Department of Geography, University of Zurich.

Svenson, O., Edland, A., and Slovic, P. (1990). Choices and judgments of incompletely described decision alternatives under time pressure. *Acta Psychologica,* vol. 75: 153-169.

Swienty, O., Reichenbacher, T., Reppermund, S., and Zihl, J. (2008). The role of relevance and cognition in attention-guiding geovisualization. *The Cartographic Journal,* vol. 45, no. 3: 227-238.

Tanner, W. P. and Swets, J. A. (1954). A decision-making theory of visual detections. *Psychological Review 61,* vol. 61: 401-409.

Terlecki, M., Newcombe, N. S., and Little, M. (2008). Durable and generalized effects on mental rotation: Gender differences in grown patterns. *Applied Cognitive Psychology,* vol. 22: 996-1013.

Thomas, L. C. and Wickens, C. D. (2006). Display dimensionality, conflict geometry, and time pressure effects on conflict detection and resolution performance using cockpit displays of traffic information. *The International Journal of Aviation Psychology,* vol. 16, no. 3: 321-342.

Torun, A. and Ulubay, A. (2005). Cartographic visualization of military maps at mid-scale. *Proceedings of the 22th International Cartography Conference (ICC 2005), A Coruña, Spain, July 9-16, 2005.*

Tufte, E. R. (1983). *The visual display of quantitative information*, Graphics Press, Cheshire, Connecticut.

Tversky, A. and Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science,* vol. 185, no. 5147: 1124-1131.

Tversky, A. and Kahneman, D. (2000). Advances in prospect theory: Cumulative representation of uncertainty. In: Kahneman, D., and Tversky, A., (eds.), *Choices, values and frames*, Cambridge University Press, Cambridge: 44-65.

Vandenberg, S. G. and Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual And Motor Skills,* vol. 47: 599-604.

Voyer, D. and Saunders, K. A. (2004). Sex differences on the mental rotation test: A factor analysis. *Acta Psychologica,* vol. 117: 79-94.

Weiss, E., Kemmler, G., Deisenhammer, E., and Fleischhacker, W. (2003). Sex differences in cognitive functions. Personality and individual differences. *Personality and Individual Differences,* vol. 35: 863-875.

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica,* vol. 41, no. 1: 67-85.

Wood, J., Kirschenbauer, S., Döllner, J., Lopes, A., and Bodum, L. (2005). Using 3D in visualization. In: Dykes, J., MacEachren, A.M., and Kraak, M.J., (eds.), *Exploring Geovisualization*, Elsevier, Oxford: 295-312.

Yates, P. (1990). *Judgment and decision making*, Prentice-Hall, Englewood Cliffs, New Jersey.

Yerkes, R. M. and Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparitive Neurology and Psychology,* vol. 18, no. 5: 459-482.

Zakay, D. (1985). Post-decision confidence and conflict experienced in a choice progress. *Acta Psychologica,* vol. 58: 75-80.

Zanola, S., Fabrikant, S. I., and Çöltekin, A. (2009). The effect of realism on the confidence in spatial data quality in stereoscopic 3D displays. *Proceedings of the 24th International Cartography Conference (ICC 2009), Santiago, Chile, Nov. 15-21, 2009.*

APPENDIX

Experiment I

Background questions

(1) **Please specify your gender** *(male / female)*
(2) **Are you wearing glasses or contact lenses?** *(Y/N)*
(3) **Have you been told by a professional that you lack depth perception?** *(Y/N)*
(4) **Have you been told by a professional that you have imperfect color vision?** *(Y/N)*
(5) **How would you rate your ability to read maps?** *(Y/N)*
(6) **How often do you pursue recreational activities that require map reading (for example, hiking, cycling, sailing or orienteering)?** *(never / occasionally / regularly)*
(7) **How would you rate your ability to read maps?** *(poor / below average / average / above average / good)*
(8) **How often do you pursue recreational activities that require map reading (for example, hiking, cycling, sailing or orienteering)?** (never/occasionally/regularly)
(9) **How much professional experience, training or college classes have you had in the following … ?** *(none / <1 year / 1-2-years / 2-5 years / >5 years)*
   GIS
   Cartography
   Computer graphics
   Fine arts, graphic design

Introductory text to NTP Scenario

The scenario for this experiment is as follows: A friend of yours is severely injured. You have to reach this person as fast as you can. You can choose from different kinds of maps to reach this person. Which map would you choose?

Below you find the six display types you can choose from. Please scroll down the page to rate them according to your preferences. Please make your choice as fast you can, if possible within 10 seconds.

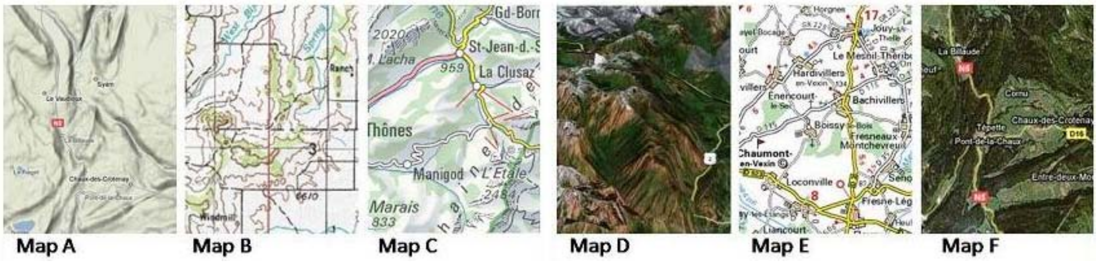
Introductory text to TP Scenario

You are on holiday and have a lot of time. You are planning an excursion for the day afterwards and can choose between different map types.

Below you find the six display types you can choose from. Please scroll down the page to rate them according to your preferences. Take as much time as you want for your choice.

Questions about map displays

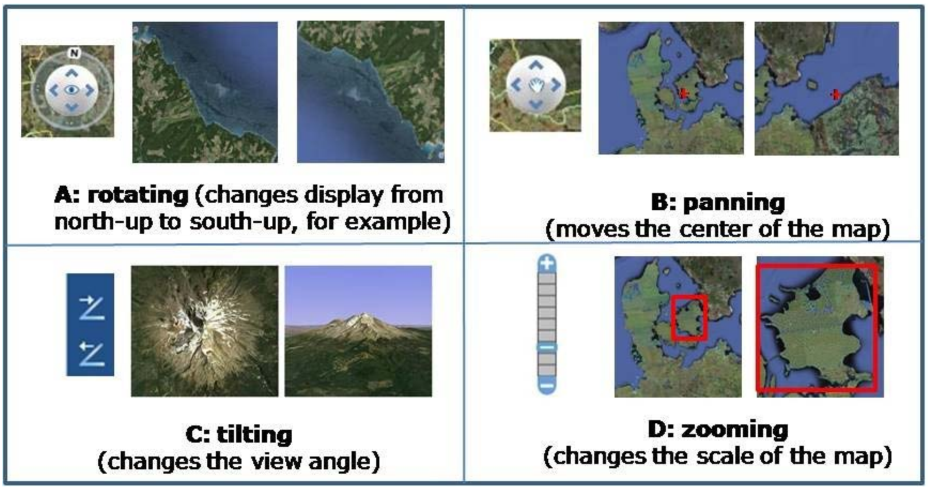
Which display type would you use for your task? Please rate the display types according to your preference from 1 (very suitable) to 5 (not suitable). *(Map A / B / C / D / E / F)*



- What did you like about your most favorite display? *(open answer)*
- What did you not like about your least favorite display? *(open answer)*

Questions about interaction tool preferences

Imagine the map you would use under time pressure was interactive.
It contains four interaction tools: rotating, tilting, zooming and panning. Below you will see four short introductory pictures about these tools.
Please rate the tools according to your preferences. Please make your choice as fast as you can, if possible within 10 seconds.



- Which interaction tool would you use? Please rate every interaction tool from 1 (very suitable) to 5 (not suitable). *(Rotating / Tilting / Zooming / Panning)*
- Why did you rate the interactivity elements the way you did? *(open answer)*

Post-test questionnaire

- Did you rate the display types differently in the two scenarios? *(Y/N)*
- If yes: Why did you choose a different rating of display types for the second scenario? *(open answer)*
- Did you rate the interaction tools differently in the two scenarios? *(Y/N)*
- If yes: Why did you choose a different rating of interactivity elements for the second scenario? *(open answer)*
- If you want, you can leave further comments on your choice and ratings here. *(open answer)*

Experiment II

Background questions

(1) **Please specify your gender** *(male / female)*
(2) **How often do you use maps in your daily life (work, studies)?**
(never / occasionally / very occasionally)
(3) **How much work experience do you have in cartography and/or Geographic Information Systems?** *(none / < 1 year / 1-2 years / 2-5 years / >5 years)*
(4) **How often do you pursue recreational activities that require map reading (for example, hiking, cycling, sailing or orienteering)?**
(never / occasionally / very occasionally)
(5) **Have you been told by a professional that you have imperfect color vision?** *(Y/N)*

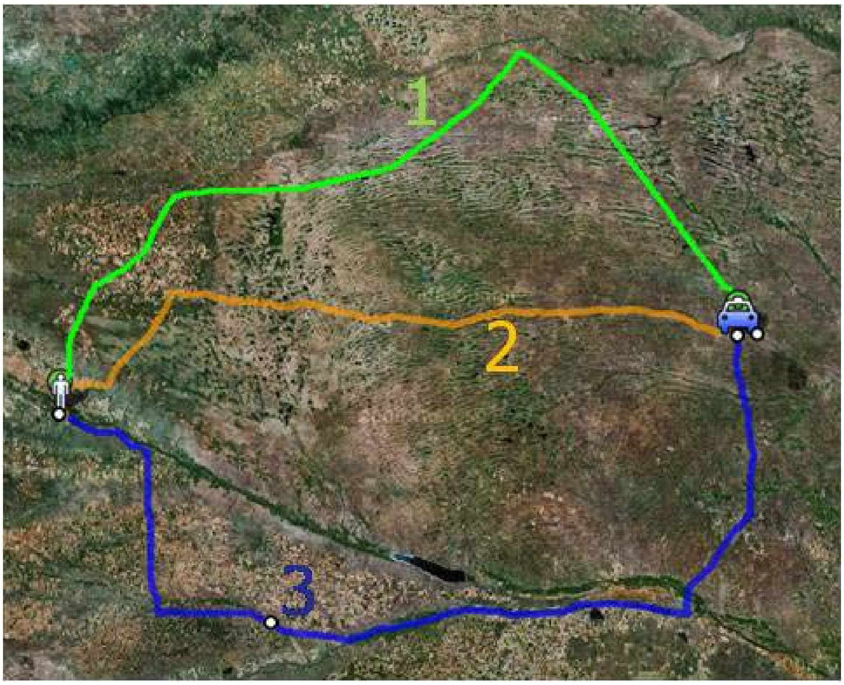
Introductory text to scenarios[15]

In this experiment, your task is to select a route under time pressure.

> **Shortest route:** Please select the shortest route (in distance, not driving time) among three different choices.
>
> **Fastest route:** Please select the route among three different choices, which is the fastest one by car from the person.

On each map, your car is your starting location, and the person symbol the end location.

For instance: In the picture below, route 2 is the shortest route.



[15] This was a between-subject experiment: Each participant either had to select only the fastest routes or only the shortest routes throughout the entire experiment.

Please note that you only have a limited amount of time for your answer – in the upper left corner and besides the map you see a window showing you how much time is remaining. You always have either 10, 20 or 30 seconds for your answer.

After every selection, you will be asked how confident you are in your decision.

Questions regarding map stimuli

- Which is the shortest (or fastest) route from the car to the person? *(Route 1 / 2 / 3)*
- How confident are you in your decision? Please rate your confidence on a scale from 1 "not confident at all" to 4 "very confident"?
  (1 not confident at all / 2 rather not confident / 3 rather confident / 4 very confident)

Experiment III

Background questions

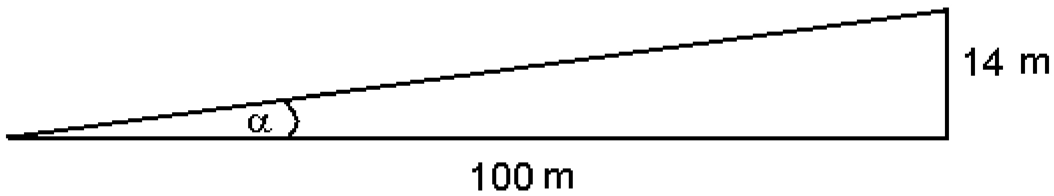
(1) **Please specify your gender** *(male / female)*
(2) **How familiar are you with topographic maps from work, studies or leisure time?** *(never / occasionally / very occasionally)*
(3) **How familiar are you with graphical representations of the third dimension (e..g, relief maps, Google Earth)?** *(none /< 1 year / 1-2 years / 2-5 years / >5 years)*
(4) **How often do you pursue recreational activities that require map reading (for example, hiking, cycling, sailing or orienteering)?** *(never / occasionally / very occasionally)*
(5) **Have you been told by a professional that you have imperfect color vision?** *(Y/N)*

Introductory text for experiment

Imagine you were a helicopter pilot. You can only land at slopes whose gradient is less than 14%.

Below you find an illustration what a gradient of 14% means: 14 meters of vertical distance (elevation difference) on 100 meters horizontal distance.

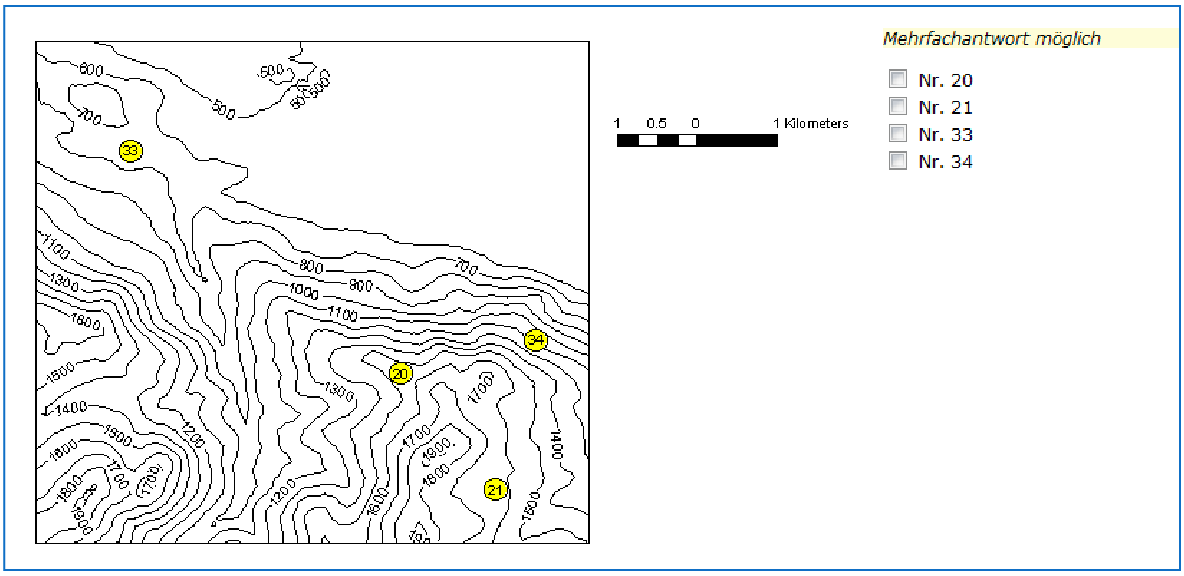
Thus, the gradient is vertical distance (elevation difference) divided by horizontal distance.



Now your task is to select points from a map, on which the gradient is below 14%. All points are numbered, and there is always at least one point at which the gradient is below 14%.

You can estimate the slope from the contour lines and the scale (at the right of the map). All maps have the same scale, and the elevation difference between the contour lines is **always 100 meters.**

On the sample map below, the points 21 and 33 are at slopes flatter than 14%. The points 20 and 34 are at slopes steeper than 14%. Thus, you would have to select 21 and 33 right of the map.

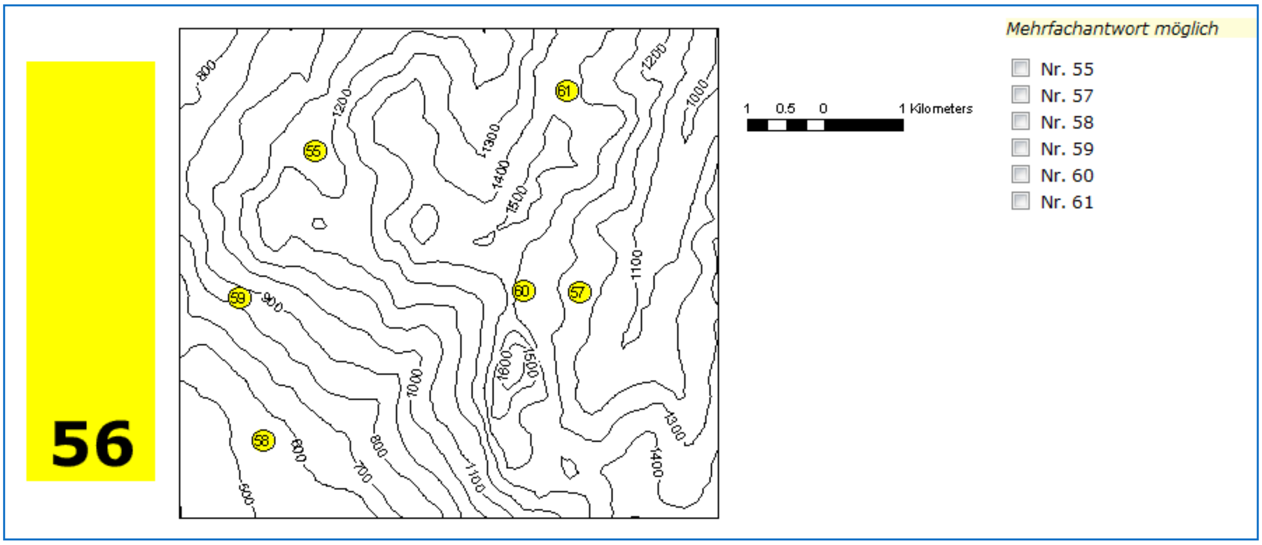


Now you can solve a sample task:

Which of these points is at a slope with less than 14% gradient?

Please note that you only have a limited amount of time for your answer – 60, 40 or 20 seconds.

The bar left of the map indicates who many seconds are remaining for your answer.



After every task, you will be asked how confident you are in your answer.

- How confident are you in your decision? Please rate your confidence on a scale from 1 "not confident at all" to 4 "very confident"?
  (1 not confident at all / 2 rather not confident / 3 rather confident / 4 very confident)

Experiment IV

Background questions (Paper & pen questionnaire)

1. **Do you use maps in your daily work life?**
   [ ] no, never [ ] yes, sometimes [ ] yes, very frequently

2. **Do you use printed paper maps in your leisure time?**
   [ ] no, never [ ] yes, sometimes [ ] yes, very frequently

   If yes, for which purpose: ____

3. **If you move with paper maps in your leisure time, do you rotate map into the direction of your movement?**
   [ ] yes, I often rotate the map [ ] no, I never rotate the map [ ] I don't know

4. **How familiar are you with virtual globes such as Google Earth?**
   [ ] not familiar at all
   [ ] rather not familiar
   [ ] rather familiar
   [ ] very familiar

5. **How familiar are you with other representations of the 3rd dimension, such as relief maps?**
   [ ] not familiar at all
   [ ] rather not familiar
   [ ] rather familiar
   [ ] very familiar

6. **Do you play video games on your computer?**

   [ ] no, never [ ] yes, sometimes [ ] yes, very frequently

   If yes, which games: ____

7. **Have you ever been told by a professional that you have imperfect color vision?**
   [ ] yes [ ] no

8. **Have you ever been told by a professional that you have imperfect stereoscopic vision?**
   [ ] yes [ ] no

CURRICULUM VITAE

JAN HENDRIK WILKENING

born January 14th, 1981, in Bückeburg, Germany

Education

2000	Graduation from high school, Kaiser-Wilhelm- und Ratsgymnasium, Hannover, Germany. Primary subjects: Latin, English, Mathematics, Politics.
2001 – 2002	Studies of Physics, Italian, and Spanish, University of Freiburg im Breisgau, Germany
2002 – 2004	Undergraduate studies in Geography (minor subjects: Geology and Economics), University of Bonn, Germany
2004 – 2005	ERASMUS scholarship at the University of Southampton, United Kingdom
2005 – 2007	Postgraduate studies in Geography (minor subjects: Geology and Economics), University of Bonn, Germany
2007	Diploma in Geography, University of Bonn, Germany. Diploma thesis: *"Entwicklung einer Vergleichsmethode für Geomarketing-Software"*, advised by Prof. Dr. Klaus Greve.
2008 – 2012	PhD studies at the Geographic Visualization and Analysis Group (GIVA), Department of Geography, University of Zurich, Switzerland. Doctoral thesis: *"The effect of time pressure on map-based decision making"*, advised by Prof. Dr. Sara Irina Fabrikant.

Publications related to PhD research

Wilkening, J. (2009): *User preferences for map-based decision making under time pressure.* In: Davies, C. (ed.): Proceedings of the Doctoral Colloquium at the 9th International Conference on Spatial Information Theory (COSIT), L'Aber Wrac'h, France, 2009, pp. 91-97.

Wilkening, J. (2010): *Map users' preferences and performance under time pressure.* In: Purves, R. and Weibel, R. (eds.): Proceedings of the 6th International Conference on Geographic Information Science (Extended Abstracts Volume), Zurich, Switzerland, September 2010.

Wilkening, J. and Fabrikant, S. I. (2011): *The effect of gender and spatial abilities on map use preferences and performance in road selection tasks.* In: Proceedings of the 25th International Cartographic Conference, Paris, France, July 2011.

Wilkening, J. and Fabrikant, S. I. (2011): *How do decision time and realism affect map-based decision making?* In: Egenhofer, M. et al. (eds.): Spatial Information Theory. 10th International Conference, COSIT 2011, Belfast, Maine, USA. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg.